UNITED STATES AIR FORCE

SUMMER RESEARCH PROGRAM -- 1995

SUMMER RESEARCH EXTENSION PROGRAM FINAL REPORTS

VOLUME 1

ARMSTRONG LABORATORY

RESEARCH & DEVELOPMENT LABORATORIES

5800 Uplander Way

Culver City, CA 90230-6608

Program Director, RDL
Gary Moore

Program Manager, AFOSR
Major David Hart

Program Manager, RDL
Scott Licoscos

Program Administrator, RDL
Gwendolyn Smith

Program Administrator
Johnetta Thompson

Submitted to:

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Bolling Air Force Base

Washington, D.C.

May 1996

REPORT DOCUMENTATION PAGE

AFRL-SR-BL-TR-00-

0696

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
May, 1996

3. REPORT T...

4. TITLE AND SUBTITLE
1995 Summer Research Program (SRP), Summer Research Extension Program (SREP), Final Report, Volume 1, Armstrong Laboratory

5. FUNDING NUMBERS
F49620-93-C-0063

6. AUTHOR(S)
Gary Moore

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Research & Development Laboratories (RDL)
5800 Uplander Way
Culver City, CA 90230-6608

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Office of Scientific Research (AFOSR)
801 N. Randolph St.
Arlington, VA 22203-1977

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
Approved for Public Release

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The United States Air Force Summer Research Program (SRP) is designed to introduce university, college, and technical institute faculty members to Air Force research. This is accomplished by the faculty members, graduate students, and high school students being selected on a nationally advertised competitive basis during the summer intersession period to perform research at Air Force Research Laboratory (AFRL) Technical Directorates and Air Force Air Logistics Centers (ALC). AFOSR also offers its research associates (faculty only) an opportunity, under the Summer Research Extension Program (SREP), to continue their AFOSR-sponsored research at their home institutions through the award of research grants. This volume consists of the SREP program background, management information, statistics, a listing of the participants, and the technical report for each participant of the SREP working at the AF Armstrong Laboratory.

14. SUBJECT TERMS
Air Force Research, Air Force, Engineering, Laboratories, Reports, Summer, Universities, Faculty, Graduate Student, High School Student

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to ***stay within the lines*** to meet ***optical scanning requirements.***

Block 1. Agency Use Only *(Leave blank)*.

Block 2. Report Date. Full publication date including day, month, and year, if available
(e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C - Contract	**PR** - Project
G - Grant	**TA** - Task
PE - Program Element	**WU** - Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. *(If known)*

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with....; Trans. of....; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

DOD - See DoDD 5230.24, "Distribution Statements on Technical Documents."
DOE - See authorities.
NASA - See Handbook NHB 2200.2.
NTIS - Leave blank.

Block 12b. Distribution Code.

DOD - Leave blank.
DOE - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.
Leave blank.
NASA - Leave blank.
NTIS -

Block 13. Abstract. Include a brief *(Maximum 200 words)* factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code *(NTIS only)*.

Blocks 17. - 19. Security Classifications. Self- explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

PREFACE

This volume is part of a five-volume set that summarizes the research of participants in the 1995 AFOSR Summer Research Extension Program (SREP). The current volume, Volume 1 of 5, presents the final reports of SREP participants at Armstrong Laboratory, Phillips Laboratory, Rome Laboratory, Wright Laboratory, Arnold Engineering Development Center, Frank J. Seiler Research Laboratory, and Wilford Hall Medical Center.

Reports presented in this volume are arranged alphabetically by author and are numbered consecutively -- e.g., 1-1, 1-2, 1-3; 2-1, 2-2, 2-3, with each series of reports preceded by a management summary. Reports in the five-volume set are organized as follows:

VOLUME	TITLE
1A	Armstrong Laboratory (part one)
1B	Armstrong Laboratory (part two)
2	Phillips Laboratory
3	Rome Laboratory
4A	Wright Laboratory (part one)
4B	Wright Laboratory (part two)
5	Arnold Engineering Development Center Frank J. Seiler Research Laboratory Wilford Hall Medical Center

1995 SREP FINAL REPORTS

Armstrong Laboratory

VOLUME 1

1995 SREP FINAL REPORTS

Armstrong Laboratory

VOLUME 1 (cont.)

1995 SREP FINAL REPORTS

Phillips Laboratory

VOLUME 2A

1995 SREP FINAL REPORTS

Phillips Laboratory

VOLUME 2B

1995 SREP FINAL REPORTS

Rome Laboratory

VOLUME 3

1995 SREP FINAL REPORTS

Wright Laboratory

VOLUME 4A

1995 SREP FINAL REPORTS

Wright Laboratory

VOLUME 4A (cont.)

1995 SREP FINAL REPORTS

Wright Laboratory

VOLUME 4B

1995 SREP FINAL REPORTS

VOLUME 5

Report #	Author's University	Report Author
	Arnold Engineering Development Center	
1	Plant-Wide Preventive Maintenance and Monitoring Vanderbilt University	Mr. Theodore Bapty Electrical Engineering AEDC
	Frank J. Seiler Research Laboratory	
1	Block Copolymers at Inorganic Solid Surfaces Colorado School of Mines, Golden, CO	Dr. John Dorgan Chemical Engineering FJSRL
2	Non-Linear Optical Properties of Polyacetylenes and Related Barry University, Miami, FL	Dr. M. A. Jungbauer Chemistry FJSRL
3	Studies of Second Harmonic Generation in Glass Waveguides Allegheny College, Meadville, PA	Dr. David Statman Physics FJSRL
	Wilford Hall Medical Center	
1	Biochemical & Cell Physiological Aspects of Hyperthermia University of Miami, Coral Gables, FL	Dr. W. Drost-Hansen Chemistry WHMC

1995 SUMMER RESEARCH EXTENSION PROGRAM (SREP) MANAGEMENT REPORT

1.0 BACKGROUND

Under the provisions of Air Force Office of Scientific Research (AFOSR) contract F49620-90-C-0076, September 1990, Research & Development Laboratories (RDL), an 8(a) contractor in Culver City, CA, manages AFOSR's Summer Research Program. This report is issued in partial fulfillment of that contract (CLIN 0003AC).

The Summer Research Extension Program (SREP) is one of four programs AFOSR manages under the Summer Research Program. The Summer Faculty Research Program (SFRP) and the Graduate Student Research Program (GSRP) place college-level research associates in Air Force research laboratories around the United States for 8 to 12 weeks of research with Air Force scientists. The High School Apprenticeship Program (HSAP) is the fourth element of the Summer Research Program, allowing promising mathematics and science students to spend two months of their summer vacations working at Air Force laboratories within commuting distance from their homes.

SFRP associates and exceptional GSRP associates are encouraged, at the end of their summer tours, to write proposals to extend their summer research during the following calendar year at their home institutions. AFOSR provides funds adequate to pay for SREP subcontracts. In addition, AFOSR has traditionally provided further funding, when available, to pay for additional SREP proposals, including those submitted by associates from Historically Black Colleges and Universities (HBCUs) and Minority Institutions (MIs). Finally, laboratories may transfer internal funds to AFOSR to fund additional SREPs. Ultimately the laboratories inform RDL of their SREP choices, RDL gets AFOSR approval, and RDL forwards a subcontract to the institution where the SREP associate is employed. The subcontract (see Appendix 1 for a sample) cites the SREP associate as the principal investigator and requires submission of a report at the end of the subcontract period.

Institutions are encouraged to share costs of the SREP research, and many do so. The most common cost-sharing arrangement is reduction in the overhead, fringes, or administrative charges institutions would normally add on to the principal investigator's or research associate's labor. Some institutions also provide other support (e.g., computer run time, administrative assistance, facilities and equipment or research assistants) at reduced or no cost.

When RDL receives the signed subcontract, we fund the effort initially by providing 90% of the subcontract amount to the institution (normally $18,000 for a $20,000 SREP). When we receive the end-of-research report, we evaluate it administratively and send a copy to the laboratory for a technical evaluation. When the laboratory notifies us the SREP report is acceptable, we release the remaining funds to the institution.

2.0 THE 1995 SREP PROGRAM

SELECTION DATA: A total of 719 faculty members (SFRP Associates) and 286 graduate students (GSRP associates) applied to participate in the 1994 Summer Research Program. From these applicants 185 SFRPs and 121 GSRPs were selected. The education level of those selected was as follows:

1994 SRP Associates, by Degree			
SFRP		GSRP	
PHD	MS	MS	BS
179	6	52	69

Of the participants in the 1994 Summer Research Program 90 percent of SFRPs and 25 percent of GSRPs submitted proposals for the SREP. Ninety proposals from SFRPs and ten from GSRPs were selected for funding, which equates to a selection rate of 54% of the SFRP proposals and of 34% for GSRP proposals.

1995 SREP: Proposals Submitted vs. Proposals Selected			
	Summer 1994 Participants	Submitted SREP Proposals	SREPs Funded
SFRP	185	167	90
GSRP	121	29	10
TOTAL	306	196	100

The funding was provided as follows:

Contractual slots funded by AFOSR	75
Laboratory funded	14
Additional funding from AFOSR	11
Total	100

Six HBCU/MI associates from the 1994 summer program submitted SREP proposals; six were selected (none were lab-funded; all were funded by additional AFOSR funds).

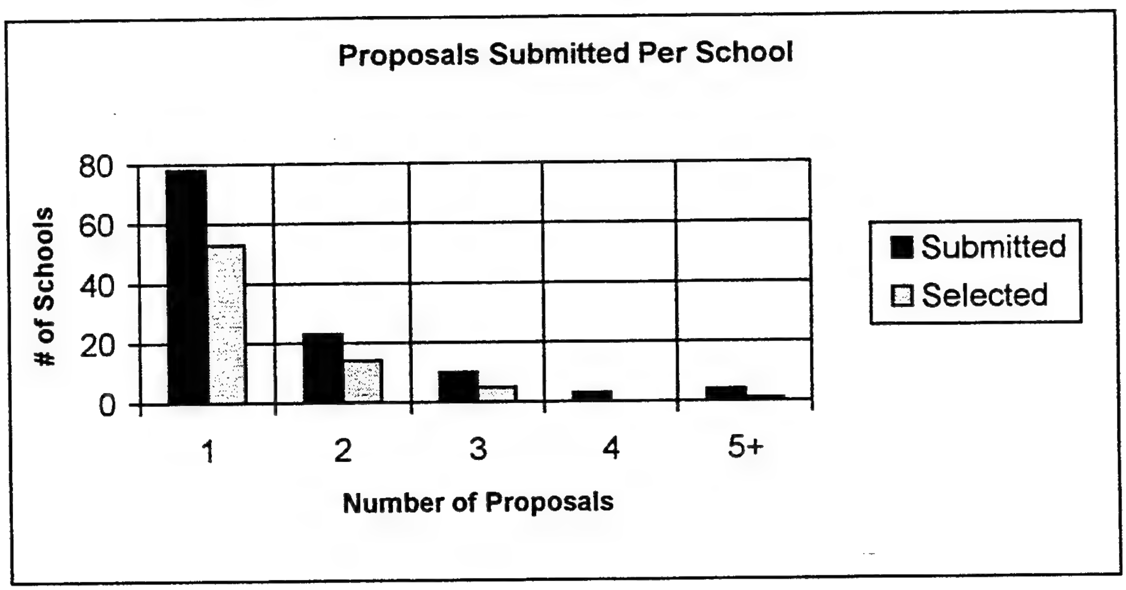
Proposals Submitted and Selected, by Laboratory		
	Applied	Selected
Armstrong Laboratory	41	19
Arnold Engineering Development Center	12	4
Frank J. Seiler Research Laboratory	6	3
Phillips Laboratory	33	19
Rome Laboratory	31	13
Wilford Hall Medical Center	2	1
Wright Laboratory	62	37
TOTAL		

Note: Phillips Laboratory funded 3 SREPs; Wright Laboratory funded 11; and AFOSR funded 11 beyond its contractual 75.

The 306 1994 Summer Research Program participants represented 135 institutions.

Institutions Represented on the 1994 SRP and 1995 SREP		
Number of schools represented in the Summer 92 Program	Number of schools represented in submitted proposals	Number of schools represented in Funded Proposals
135	118	73

Forty schools had more than one participant submitting proposals.



The selection rate for the 78 schools submitting 1 proposal (68%) was better than those submitting 2 proposals (61%), 3 proposals (50%), 4 proposals (0%) or 5+ proposals (25%). The 4 schools that submitted 5+ proposals accounted for 30 (15%) of the 196 proposals submitted.

Of the 196 proposals submitted, 159 offered institution cost sharing. Of the funded proposals which offered cost sharing, the minimum cost share was $1000.00, the maximum was $68,000.00 with an average cost share of $12,016.00.

Proposals and Institution Cost Sharing		
	Proposals Submitted	Proposals Funded
With cost sharing	159	82
Without cost sharing	37	18
Total	196	100

The SREP participants were residents of 41 different states. Number of states represented at each laboratory were:

States Represented, by Proposals Submitted/Selected per Laboratory		
	Proposals Submitted	Proposals Funded
Armstrong Laboratory	21	13
Arnold Engineering Development Center	5	2
Frank J. Seiler Research Laboratory	5	3
Phillips Laboratory	16	14
Rome Laboratory	14	7
Wilford Hall Medical Center	2	1
Wright Laboratory	24	20

Eleven of the 1995 SREP Principal Investigators also participated in the 1994 SREP.

<u>ADMINISTRATIVE EVALUATION</u>: The administrative quality of the SREP associates' final reports was satisfactory. Most complied with the formatting and other instructions provided to them by RDL. Ninety seven final reports and two interim reports have been received and are included in this report. The subcontracts were funded by $1,991,623.00 of Air Force money. Institution cost sharing totaled $985,353.00.

TECHNICAL EVALUATION: The form used for the technical evaluation is provided as Appendix 2. ninety-two evaluation reports were received. Participants by laboratory versus evaluations submitted is shown below:

	Participants	Evaluations	Percent
Armstrong Laboratory	23[1]	20	95.2
Arnold Engineering Development Center	4	4	100
Frank J. Seiler Research Laboratory	3	3	100
Phillips Laboratory	19[2]	18	100
Rome Laboratory	13	13	100
Wilford Hall Medical Center	1	1	100
Wright Laboratory	37	34	91.9
Total			

Notes:

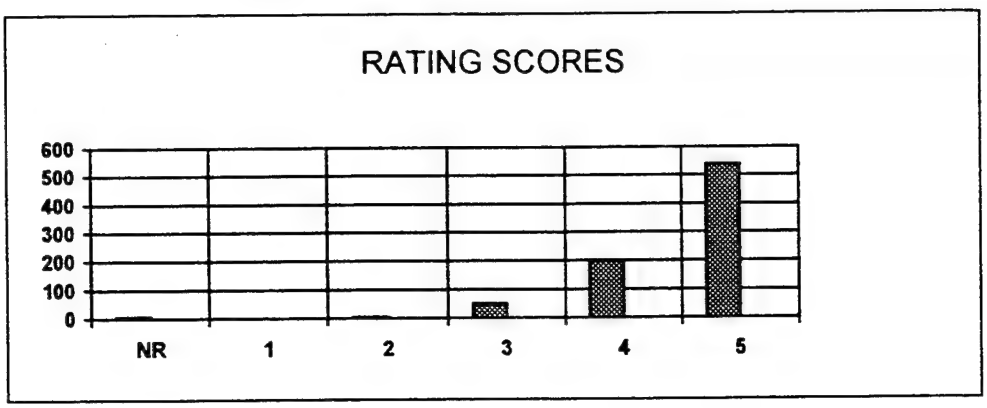
1: Research on two of the final reports was incomplete as of press time so there aren't any technical evaluations on them to process, yet. Percent complete is based upon 20/21=95.2%

2: One technical evaluation was not completed because one of the final reports was incomplete as of press time. Percent complete is based upon 18/18=100%

3: See notes 1 and 2 above. Percent complete is based upon 93/97=95.9%

The number of evaluations submitted for the 1995 SREP (95.9%) shows a marked improvement over the 1994 SREP submittals (65%).

PROGRAM EVALUATION: Each laboratory focal point evaluated ten areas (see Appendix 2) with a rating from one (lowest) to five (highest). The distribution of ratings was as follows:



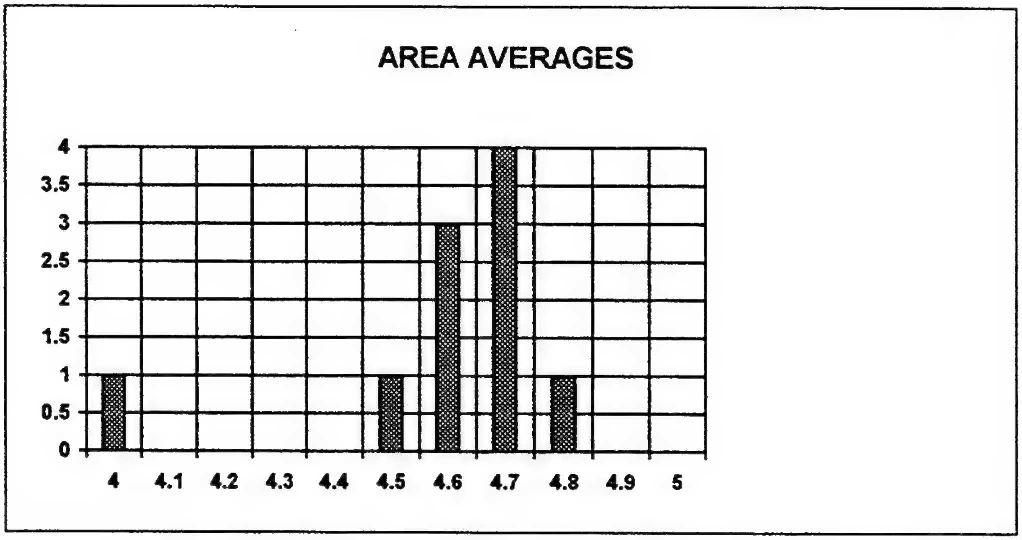
Rating	Not Rated	1	2	3	4	5
# Responses	7	1	7	62 (6%)	226 (25%)	617 (67%)

The 8 low ratings (one 1 and seven 2's) were for question 5 (one 2) "The USAF should continue to pursue the research in this SREP report" and question 10 (one 1 and six 2's) "The

one-year period for complete SREP research is about right", in addition over 30% of the threes (20 of 62) were for question ten. The average rating by question was:

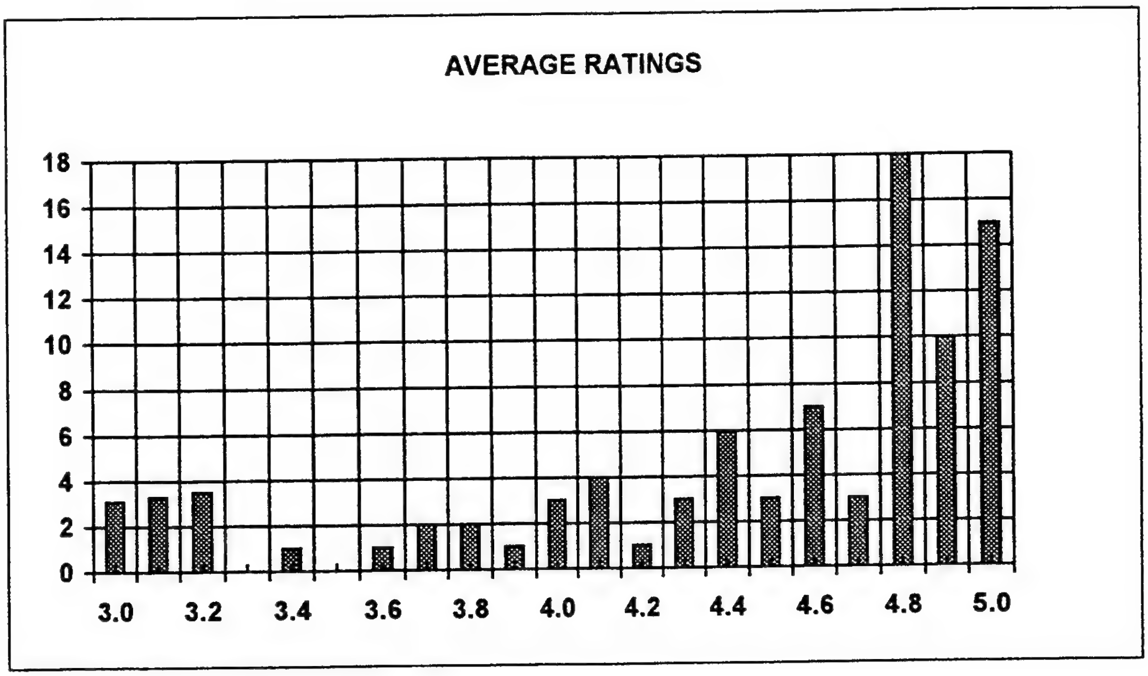
Question	1	2	3	4	5	6	7	8	9	10
Average	4.6	4.6	4.7	4.7	4.6	4.7	4.8	4.5	4.6	4.0

The distribution of the averages was:



Area 10 "the one-year period for complete SREP research is about right" had the lowest average rating (4.1). The overall average across all factors was 4.6 with a small sample standard deviation of 0.2. The average rating for area 10 (4.1) is approximately three sigma lower than the overall average (4.6) indicating that a significant number of the evaluators feel that a period of other than one year should be available for complete SREP research.

The average ratings ranged from 3.4 to 5.0. The overall average for those reports that were evaluated was 4.6. Since the distribution of the ratings is not a normal distribution the average of 4.6 is misleading. In fact over half of the reports received an average rating of 4.8 or higher. The distribution of the average report ratings is as shown:



It is clear from the high ratings that the laboratories place a high value on AFOSR's Summer Research Extension Programs.

3.0 SUBCONTRACTS SUMMARY

Table 1 provides a summary of the SREP subcontracts. The individual reports are published in volumes as shown:

Laboratory	Volume
Armstrong Laboratory	1A, 1B
Arnold Engineering Development Center	5
Frank J. Seiler Research Laboratory	5
Phillips Laboratory	2
Rome Laboratory	3
Wilford Hall Medical Center	5
Wright Laboratory	4A, 4B

1995 SREP SUB-CONTRACT DATA

Report Author Author's University	Author's Degree	Sponsoring Lab	Performance Period		Contract Amount	Univ. Cost Share
Anderson , James Analytical Chemistry University of Georgia, Athens, GA	PhD 95-0807	AL/EQ	01/01/95	12/31/95	$25000.00	$1826.00
		Determination of the Redox Capacity of Soil Sediment and Prediction of Pollutant				
Ashrafiuon , Hashem Mechanical Engineering Villanova University, Villanova, PA	PhD 95-0800	AL/CF	01/01/95	12/31/95	$25000.00	$19528.00
		Finite Element Modeling of the Human Neck and Its Validation for the ATB Model				
Burke , Michael Tulane University Tulane University, New Orleans, LA	PhD 95-0811	AL/HR	01/01/95	09/30/95	$25000.00	$1818.00
		An Examination of the Validity of the New Air Force ASVAB Composites				
Edwards , Paul Chemistry Edinboro Univ of Pennsylvania, Edinboro, PA	PhD 95-0808	AL/EQ	01/01/95	12/31/95	$25000.00	$5000.00
		Fuel Identification by Neural Networks Analysis of the Response of Vapor Sensiti				
Gerstman , Bernard Physics Florida International Universi, Miami, FL	PhD 95-0815	AL/OE	01/01/95	12/31/95	$24289.00	$2874.00
		A Comparison of Multistep vs Singlestep Arrhenius Integral Models for Describing				
Graetz , Kenneth Department of Psychology University of Dayton, Dayton, OH	PhD 95-0812	AL/HR	01/01/95	12/31/95	$25000.00	$0.00
		Effects of Mental Workload and Electronic Support on Negotiation Performance				
Gupta , Pushpa Mathematics University of Maine, Orono, ME	PhD 95-0802	AL/AO	01/01/95	12/31/95	$25000.00	$2859.00
		Regression to the Mean in Half Life Studies				
Koch , Manfred Geophysics Florida State University, Tallahassee, FL	PhD 95-0809	AL/EQ	12/01/94	04/30/95	$25000.00	$0.00
		Application of the MT3D Solute Transport Model to the Made-2 Site: Calibration				
Novotny , Mark Supercomputer Comp Res. I Florida State University, Tallahassee, FL	PhD 95-0810	AL/EQ	01/01/95	12/31/95	$25000.00	$0.00
		Computer Calculations of Gas-Phase Reaction Rate Constants				
Nurre , Joseph Mechanical Engineering Ohio University, Athens, OH	PhD 95-0804	AL/CF	01/01/95	12/31/95	$25000.00	$20550.00
		Surface Fitting Three Dimensional Human Head Scan Data				
Piepmeier , Edward Pharmaceutics University of South Carolina, Columbia, SC	PhD 95-0801	AL/AO	01/01/95	12/31/95	$25000.00	$11740.00
		The Effects of Hyperbaric Oxygenation on Metabolism of Drugs and Other Xenobioti				
Quinones , Miguel Psychology Rice University, Houston, TX	PhD 95-0813	AL/HR	01/01/95	12/31/95	$25000.00	$4000.00
		Maintaining Skills After Training: The Role of Opportunity to Perform Trained T				
Riccio , Gary Psychology Univ of IL Urbana-Champaign, Urbana, IL	PhD 95-0806	AL/CF	01/01/95	05/31/95	$22931.00	$0.00
		Nonlinear Transcutaneous Electrical Stimulation of the Vestibular System				
Shebilske , Wayne Dept of Psychology Texas A&M University, College Station, TX	PhD 95-0814	AL/HR	01/01/95	12/31/95	$25000.00	$5614.00
		Cognitive Factors in Distr Training Effects During Acquisition of Complex Skills				

1995 SREP SUB-CONTRACT DATA

Report Author / Author's University	Author's Degree	Sponsoring Lab	Performance Period		Contract Amount	Univ. Cost Share
Weisenberger , Janet Dept of Speech & Hearing Ohio State University, Columbus, OH	PhD 95-0805	AL/CF	01/01/95	12/31/95	$25000.00	$12234.00
		Tactile Feedback for Simulation of Object Shape and Textural Information in Hapt				
Hughes , Rod Psychology Oregon Health Sciences University, Portland, OR	MA 95-0803	AL/CF	01/01/95	12/31/95	$25000.00	$0.00
		Melatonin Induced Prophylactic Sleep as a Countermeasure for Sleep Deprivation				
Bapty , Theodore Electrical Engineering Vanderbilt University, Nashville, TN	MS 95-0848	AEDC/E	01/01/95	12/31/95	$24979.00	$0.00
		Plant-Wide Preventive Maintenance & Monitoring				
Dorgan , John Chemical Engineering Colorado School of Mines, Golden, CO	PhD 95-0834	FJSRL/F	01/01/95	12/31/95	$25000.00	$0.00
		Block Copolymers at Inorganic Solid Surfaces				
Jungbauer , Mary Ann Chemistry Barry University, Miami, FL	PhD 95-0836	FJSRL/F	01/01/95	12/31/95	$25000.00	$24714.00
		Non-Linear Optical Properties of Polyacetylenes and Related Substituted Compound				
Statman , David Physics Allegheny College, Meadville, PA	PhD 95-0835	FJSRL/F	01/01/95	12/31/95	$25000.00	$6500.00
		Studies of Second Harmonic Generation in Glass Waveguides				
, Krishnaswamy Aeronautics University of Houston, Houston, TX	PhD 95-0818	PL/RK	01/01/95	12/31/95	$24993.00	$8969.00
		Mixed-Mode Fracture of Solid Propellants				
Ashgriz , Nasser Mechanical Engineering SUNY-Buffalo, Buffalo, NY	PhD 95-0816	PL/RK	01/01/95	12/31/95	$25000.00	$22329.00
		Effects of the Jet Characteristics on the Atomization and Mixing in A Pair of Im				
Bellem , Raymond Computer Science Embry-Riddle Aeronautical Univ, Prescott, AZ	PhD 95-0817	PL/VT	12/01/94	11/30/95	$20000.00	$8293.00
		Experimental Studies of the Effects of Ionizing Radiation on Commerically Proces				
Brzosko , Jan Nuclear Physics Stevens Institute of Tech, Hoboken, NJ	PhD 95-0828	PL/WS	11/01/94	02/01/95	$24943.00	$0.00
		Neutron Diagnostics for Pulsed Plasmas of Compact Toroid - Marauder Type				
Damodaran , Meledath Math & Computer Science University of Houston-Victoria, Victoria, TX	PhD 95-0831	PL/LI	01/01/95	12/31/94	$24989.00	$9850.00
		Parallel Computation of Zernike Aberration Coefficients for Optical Aber Correct				
DeLyser , Ronald Electrical Engineering University of Denver, Denver, CO	PhD 95-0877	PL/WS	01/01/95	12/31/95	$25000.00	$46066.00
		Quality Factor Evaluation of Complex Cavities				
Diels , Jean-Claude Physics University of New Mexico, Albuquerque, NM	PhD 95-0819	PL/LI	01/01/95	12/31/95	$25000.00	$0.00
		Unidirectional Ring Lasers and Laseer Gyros with Multiple Quantum Well Gain Medi				
Henson , James Electrical Engineering University of Nevada, Reno, NV	PhD 95-0820	PL/WS	01/01/95	12/31/95	$25000.00	$0.00
		Automatic Feature Extraction and Assessment of Wideband Range-Doppler Imagery of				
Kaiser , Gerald Physics University of Mass/Lowell, Lowell, MA	PhD 95-0821	PL/GP	01/01/95	12/31/95	$25000.00	$5041.00
		Multiresolution Analysis with Physical Wavelets				

1995 SREP SUB-CONTRACT DATA

Report Author Author's University	Author's Degree	Sponsoring Lab	Performance Period	Contract Amount	Univ. Cost Share
Kowalak , Albert Chemistry 95-0822 University of Massachusetts/Lo, Lowell, MA	PhD	PL/GP	01/01/95 12/31/95	$24996.00	$4038.00
		The Synthesis and Chemistry of Peroxonitrites and Peroxonitrous Acid			
Malloy , Kevin Electrical Engineering 95-0829 University of New Mexico, Albuquerque, NM	PhD	PL/VT	01/01/95 12/31/95	$24999.00	$0.00
		Temperature & Pressure Dependence of the Band Gaps & Band Offsets			
Prasad , Sudhakar Physics 95-0823 University of New Mexico, Albuquerque, NM	PhD	PL/LI	01/01/95 12/31/95	$25000.00	$11047.00
		Theoretical Studies of the Performance of Novel Fiber-Coupled Imaging Interferom			
Purtill , Mark Mathematics 95-0824 Texas A&M Univ-Kingsville, Kingsville, TX	PhD	PL/WS	01/01/95 12/31/95	$25000.00	$100.00
		Static and Dynamic Graph Embedding for Parallel Programming			
Rudolph , Wolfgang Physics 95-0833 University of New Mexico, Albuquerque, NM	PhD	PL/LI	01/01/95 12/31/95	$24982.00	$6000.00
		Ultrafast Process and Modulation in Iodine Lasers			
Stone , Alexander Mathematics & Statistics 95-0827 University of New Mexico, Alburquerque, NM	PhD	PL/WS	01/01/95 12/31/95	$24969.00	$0.00
		Impedance Matching And Reflection Minimization For Transient EM Pulses Through D			
Swenson , Charles Dept of Electrical Engnr 95-0826 Utah State University, Logan, UT	PhD	PL/VT	01/01/95 12/31/95	$25000.00	$25000.00
		Low Power Retromodulator based Optical Transceiver for Satellite Communications			
Lipp , John Electrical Engineering 95-0832 Michigan Technological Univ, Houghton, MI	MS	PL/LI	01/01/95 12/31/95	$24340.00	$15200.00
		Improved Methods of Tilt Measurement for Extended Images in the Presence of Atmo			
Petroski , Janet Chemistry 95-0830 Cal State Univ/Northridge, Northridge, CA	BA	PL/RK	10/01/94 12/31/94	$4279.00	$0.00
		Thermoluminescence of Simple Species in Molecular Hydrogen Matrices			
Salasovich , Richard Mechanical Engineering 95-0825 University of Cincinnati, Cincinnati, OH	MS	PL/VT	01/01/95 12/31/95	$25000.00	$4094.00
		Design, Fabrication, Intelligent Cure, Testing, and Flight Qualification of an A			
Aalo , Valentine Dept of Electrical Engr 95-0837 Florida Atlantic University, Boca Raton, FL	PhD	RL/C3	01/01/95 12/31/95	$25000.00	$13120.00
		Performance Study of an ATM/Satellite Network			
Amin , Moeness Electrical Engineering 95-0838 Villanova University, Villanova, PA	PhD	RL/C3	01/01/95 12/31/95	$25000.00	$34000.00
		Interference Excision in Spread Spectrum Communication Systems Using Time-Freque			
Benjamin , David Computer Science 95-0839 Oklahoma State University, Stillwater, OK	PhD	RL/C3	01/01/95 12/31/95	$24970.00	$0.00
		Designing Software by Decomposition using KIDS			
Choudhury , Ajit Engineering 95-0840 Howard University, Washington, DC	PhD	RL/OC	11/30/94 10/31/95	$25000.00	$0.00
		Detection Performance of Over Resolved Targets with Non-Uniform and Non-GAussian			
Harackiewicz , Frances Electrical Engineering 95-0841 So. Illinois Univ-Carbondale, Carbondale, IL	PhD	RL/ER	01/01/95 12/31/95	$23750.00	$29372.00
		Computer-Aided-Design Program for Solderless Coupling Between Microstrip and Str			

1995 SREP SUB-CONTRACT DATA

Report Author Author's University	Author's Degree	Sponsoring Lab	Performance Period	Contract Amount	Univ. Cost Share
Losiewicz , Beth Psycholinguistics Colorado State University, Fort Collins, CO	PhD 95-0842	RL/IR	01/01/95 12/31/95	$25000.00	$4850.00
		Spanish Dialect Identification Project			
Musavi , Mohamad University of Maine, Orono, ME	PhD 95-0843	RL/IR	01/01/95 12/31/95	$25000.00	$12473.00
		Automatic Image Registration Using Digital Terrain Elevation Data			
Norgard , John Elec & Comp Engineering Univ of Colorado-Colorado Sprg, Colorado	PhD 95-0844	RL/ER	01/01/95 12/31/95	$25000.00	$2500.00
		Infrared Images of Electromagnetic Fields			
Richardson , Dean Photonics SUNY Institute of Technology, Utica, NY	PhD 95-0845	RL/OC	01/01/95 12/31/95	$25000.00	$15000.00
		Femtosecond Pump-Probe Spectroscopy System			
Ryder, Jr. , Daniel Chemical Engineering Tufts University, Medford, MA	PhD 95-0846	RL/ER	01/01/95 12/31/95	$25000.00	$0.00
		Synthesis and Properties B-Diketonate-Modified Heterobimetallic Alkoxides			
Zhang , Xi-Cheng Physics Rensselaer Polytechnic Institu, Troy, NY	PhD 95-0847	RL/ER	01/01/95 12/31/95	$25000.00	$0.00
		Optoelectronic Study of Seniconductor Surfaces and Interfaces			
Drost-Hansen , Walter Chemistry University of Miami, Coral Gables, FL	PhD 95-0875	WHMC/	01/01/95 12/31/95	$25000.00	$8525.00
		Biochemical & Cell Physiological Aspects of Hyperthermia			
Baginski , Michael Electrical Engineering Auburn University, Auburn, AL	PhD 95-0869	WL/MN	01/01/95 12/31/95	$24995.00	$10098.00
		An Investigation of the Heating and Temperature Distribution in Electrically Exc			
Berdichevsky , Victor Aerospace Engineering Wayne State University, Detroit, MI	PhD 95-0849	WL/FI	01/01/95 12/31/95	$25000.00	$0.00
		Micromechanics of Creep in Metals and Ceramics at High Temperature			
Buckner , Steven Chemistry Colullmbus College, Columbus, GA	PhD 95-0850	WL/PO	01/01/95 12/31/95	$24900.00	$8500.00
		Development of a Fluorescenece-Based Chemical Sensor for Simultaneous Oxygen Qua			
Carroll , James Electrical Engineering Clarkson University, Potsdam, NY	PhD 95-0881	WL/PO	01/01/95 12/31/95	$24944.00	$38964.00
		Development of HIgh-Performance Active Dynamometer System for Machines and Drive			
Choate , David Mathematics Transylvania University, Lexington, KY	PhD 95-0851	WL/AA	01/01/95 12/31/95	$24993.00	$8637.00
		SOLVING z(t)=ln{/A[cos(w1t)]+B[sin(w2t)]+C}			
Clarson , Stephen Materials Sci & Eng University of Cincinnati, Cincinnati, OH	PhD 95-0852	WL/ML	12/01/94 11/30/95	$25000.00	$15000.00
		Synthesis, Processing and Characterization of Nonlinear Optical Polymer Thin Fil			
Cone , Milton Comp Science & Elec Eng Embry-Riddel Aeronautical Univ, Prescott, AZ	PhD 95-0853	WL/AA	01/01/95 12/31/95	$25000.00	$11247.00
		An Investigation of Planning and Scheduling Algorithms for Sensor Management			
Courter , Robert Mechanical Engineering Louisiana State University, Baton Rouge, LA	PhD 95-0854	WL/MN	01/01/95 12/31/95	$25000.00	$3729.00
		A Study to Determine Wave Gun Firing Cycles for High Performance Model Launches			

1995 SREP SUB-CONTRACT DATA

Report Author Author's University	Author's Degree	Sponsoring Lab	Performance Period		Contract Amount	Univ. Cost Share
Dominic , Vincent Electro Optics Program University of Dayton, Dayton, OH	PhD 95-0868	WL/ML Characterization of Electro-Optic Polymers	01/01/95	12/31/95	$25000.00	$12029.00
Fadel , Georges Dept of Mechanical Engnr Clemson University, Clemson, SC	PhD 95-0855	WL/MT A Methodology for Affordability in the Design Process	01/01/95	12/31/95	$25000.00	$8645.00
Gould , Richard Mechanical Engineering North Carolina State Univ, Raleigh, NC	PhD 95-0856	WL/PO Data Reduction and Analysis for laser Doppler Velocimetry	01/01/95	12/31/95	$24998.00	$9783.00
Hardie , Russell Electrical Engineering Univsity of Dayton, Dayton, OH	PhD 95-0882	WL/AA Hyperspectral Target Identification Using Bomen Spectometer Data	01/01/95	12/31/95	$24999.00	$7415.00
Hodel , Alan Electrical Engineering Auburn University, Auburn, AL	PhD 95-0870	WL/MN Robust Falut Tolerant Control: Fault Detection and Classification	01/01/95	12/31/95	$24990.00	$9291.00
Janus , Jonathan Aerospace Engineering Mississippi State University, Mississippi State,	PhD 95-0871	WL/MN Multidimensional Algorithm Development & Analysis	01/01/95	12/31/95	$25000.00	$7143.00
Jasiuk , Iwona Dept of Materials Science Michigan State University, East Lansing, MI	PhD 95-0857	WL/ML Characterization of Interfaces in Metal-Matrix Composites	01/01/95	12/31/95	$25000.00	$0.00
Jouny , Ismail Electrical Engineering Lafayette College, Easton, PA	PhD 95-0880	WL/AA TSI Mitigation: A Mountaintop Database Study	01/01/95	12/31/95	$24300.00	$5200.00
Li , Jian Electrical Engineering University of Florida, Gainesville, FL	PhD 95-0859	WL/AA Comparative Study and Performance Analysis of High Resolution SAR Imaging Techni	10/10/95	12/31/95	$25000.00	$4000.00
Lin , Chun-Shin Electrical Engineering University of Missouri-Columbi, Columbia, MO	PhD 95-0883	WL/FI Prediction of Missile Trajectory	01/01/95	12/31/95	$25000.00	$2057.00
Lin , Paul Mechanical Engineering Cleveland State University, Cleveland, OH	PhD 95-0860	WL/FI Three Dimensional Deformation Comparison Between Bias and Radial Aircraft Tires	01/01/95	12/31/95	$25000.00	$6886.00
Liou , Juin Electrical Engineering University of Central Florida, Orlando, FL	PhD 95-0876	WL/EL Investigation of AlGaAs/GaAs Heterojunction Bipolar Transister Reliability Based	01/01/95	12/31/95	$25000.00	$11040.00
Nandhakumar , Nagaraj Electrical Engineering University of Virginia, Charlottesville, VA	PhD 95-0861	WL/AA Thermophysical Invariants fro, LWIR Imagery for ATR	01/01/95	12/31/95	$24979.00	$4500.00
Pasala , Krishna Dept of Electrical Engr University of Dayton, Dayton, OH	PhD 95-0879	WL/AA Effect of Electromagmetic Enviornment on Array Signal Processing	01/01/95	12/31/95	$25000.00	$1078.00
Perkowski , Marek Dept of Electrical Engnr Portland State University, Portland, OR	PhD 95-0878	WL/AA Functional Decomposition of Binary,Multiple-Valued, & Fuzzy Logic	01/01/95	09/15/95	$24947.00	$18319.00

1995 SREP SUB-CONTRACT DATA

Report Author Author's University	Author's Degree	Sponsoring Lab	Performance Period	Contract Amount	Univ. Cost Share
Reeves , Stanley Dept of Electrical Engnr Auburn University, Auburn, AL	PhD 95-0862	WL/MN Superresolution of Passive Millimeter-Wave Imaging	01/01/95 12/31/95	$25000.00	$0.00
Rule , William Engineering Mechanics University of Alabama, Tuscaloosa, AL	PhD 95-0872	WL/MN Development of a Penetrator Optimizer	01/01/95 12/31/95	$24968.00	$14576.00
Schauer , John Mech & Aerosp Eng University of Dayton, Dayton, OH	PhD 95-0873	WL/PO Heat Transfer for Turbine Blade Film Cooling with Free Stream Turbulence - Measu	11/01/94 11/30/95	$25000.00	$7428.00
Schwartz , Carla Electrical Engineering University of Florida, Gainesville, FL	PhD 95-0863	WL/FI Neural Network Identification and Control in Metal Forging	01/01/95 12/31/95	$25000.00	$0.00
Simon , Terrence Dept of Mechanical Engineering University of Minnesota, Minneapolis, MN	PhD 95-0864	WL/PO Documentation of Separating and Separated Boundary Layer Flow, for Application	01/01/95 12/31/95	$24966.00	$3996.00
Skowronski , Marek Solid State Physics Carnegie Melon University, Pittsburgh, PA	PhD 95-0865	WL/EL Transmission Electron Microscopy of Semiconductor Heterojunctions	01/01/95 12/31/95	$25000.00	$6829.00
Thirunarayan , Krishnaprasad Computer Science Wright State University, Dayton, OH	PhD 95-0866	WL/EL VHDL-93 Parser in SWI-PROLOG: A Basis for Design Query System	01/01/95 12/31/95	$25000.00	$2816.00
Trelease , Robert Dept of Anatomy & Cell Bi University of California, Los Angeles, CA	PhD 95-0867	WL/ML Development of Qualitative Process Control Discovery Systems for Polymar Composi	12/01/94 12/01/95	$25000.00	$0.00
Tsai , Chi-Tay Engineering Mechanics Florida Atlantic University, Boca Raton, FL	PhD 95-0874	WL/MN Improved Algorithm Development of Massively Parallel Epic Hydrocode in Cray T3D	01/01/95 12/31/95	$24980.00	$0.00
Lewis , John Materials Science Engrng University of Kentucky, Lexington, KY	MS 95-0858	WL/MN The Characterization of the Mechanical Properties of Materials in a Biaxial Stre	01/01/95 12/31/95	$25000.00	$13833.00

APPENDIX 1:

SAMPLE SREP SUBCONTRACT

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
1995 SUMMER RESEARCH EXTENSION PROGRAM
SUBCONTRACT 95-0837

BETWEEN

Research & Development Laboratories
5800 Uplander Way
Culver City, CA 90230-6608

AND

Florida Atlantic University
Department of Electrical Engineering
Boca Raton, FL 33431

REFERENCE: Summer Research Extension Program Proposal 95-0837
Start Date: 01-01-95 End Date: 12-31-95
Proposal Amount: $25,000.00

(1) PRINCIPAL INVESTIGATOR: Dr. Valentine A. Aalo
Department of Electrical Engineering
Florida Atlantic University
Boca Raton, FL 33431

(2) UNITED STATES AFOSR CONTRACT NUMBER: F49620-93-C-0063

(3) CATALOG OF FEDERAL DOMESTIC ASSISTANCE NUMBER (CFDA):12.800
PROJECT TITLE: AIR FORCE DEFENSE RESEARCH SOURCES PROGRAM

(4) ATTACHMENT 1 REPORT OF INVENTIONS AND SUBCONTRAACT
2 CONTRACT CLAUSES
3 FINAL REPORT INSTRUCTIONS

*****SIGN SREP SUBCONTRACT AND RETURN TO RDL*****

1. BACKGROUND: Research & Development Laboratories (RDL) is under contract (F49620-93-C-0063) to the United States Air Force to administer the Summer Research Program (SRP), sponsored by the Air Force Office of Scientific Research (AFOSR), Bolling Air Force Base, D.C. Under the SRP, a selected number of college faculty members and graduate students spend part of the summer conducting research in Air Force laboratories. After completion of the summer tour participants may submit, through their home institutions, proposals for follow-on research. The follow-on research is known as the Summer Research Extension Program (SREP). Approximately 61 SREP proposals annually will be selected by the Air Force for funding of up to $25,000; shared funding by the academic institution is encouraged. SREP efforts selected for funding are administered by RDL through subcontracts with the institutions. This subcontract represents an agreement between RDL and the institution herein designated in Section 5 below.

2. RDL PAYMENTS: RDL will provide the following payments to SREP institutions:

   - 80 percent of the negotiated SREP dollar amount at the start of the SREP research period.
   - The remainder of the funds within 30 days after receipt at RDL of the acceptable written final report for the SREP research.

3. INSTITUTION'S RESPONSIBILITIES: As a subcontractor to RDL, the institution designated on the title page will:

a. Assure that the research performed and the resources utilized adhere to those defined in the SREP proposal.

b. Provide the level and amounts of institutional support specified in the SREP proposal..

c. Notify RDL as soon as possible, but not later than 30 days, of any changes in 3a or 3b above, or any change to the assignment or amount of participation of the Principal Investigator designated on the title page.

d. Assure that the research is completed and the final report is delivered to RDL not later than twelve months from the effective date of this subcontract, but no later than December 31, 1998. The effective date of the subcontract is one week after the date that the institution's contracting representative signs this subcontract, but no later than January 15, 1998.

e. Assure that the final report is submitted in accordance with Attachment 3.

f. Agree that any release of information relating to this subcontract (news releases, articles, manuscripts, brochures, advertisements, still and motion pictures, speeches, trade associations meetings, symposia, etc.) will include a statement that the project or effort depicted was or is sponsored by: Air Force Office of Scientific Research, Bolling AFB, D.C.

g. Notify RDL of inventions or patents claimed as the result of this research as specified in Attachment 1.

h. RDL is required by the prime contract to flow down patent rights and technical data requirements to this subcontract. Attachment 2 to this subcontract

contains a list of contract clauses incorporated by reference in the prime contract.

4. All notices to RDL shall be addressed to:

RDL AFOSR Program Office
5800 Uplander Way
Culver City, CA 90230-6609

5. By their signatures below, the parties agree to provisions of this subcontract.

Abe Sopher
RDL Contracts Manager

Signature of Institution Contracting Official

Typed/Printed Name

Date

Title

Institution

Date/Phone

ATTACHMENT 2
CONTRACT CLAUSES

This contract incorporates by reference the following clauses of the Federal Acquisition Regulations (FAR), with the same force and effect as if they were given in full text. Upon request, the Contracting Officer or RDL will make their full text available (FAR 52.252-2).

FAR CLAUSES	TITLE AND DATE
52.202-1	DEFINITIONS
52.203-3	GRATUITIES
52.203-5	COVENANT AGAINST CONTINGENT FEES
52.203-6	RESTRICTIONS ON SUBCONTRACTOR SALES TO THE GOVERNMENT
52.203-7	ANTI-KICKBACK PROCEDURES
52.203-8	CANCELLATION, RECISSION, AND RECOVERY OF FUNDS FOR ILLEGAL OR IMPROPER ACTIVITY
52.203-10	PRICE OR FEE ADJUSTMENT FOR ILLEGAL OR IMPROPER ACTIVITY
52.203-12	LIMITATION ON PAYMENTS TO INFLUENCE CERTAIN FEDERAL TRANSACTIONS
52.204-2	SECURITY REQUIREMENTS
52.209-6	PROTECTING THE GOVERNMENT'S INTEREST WHEN SUBCONTRACTING WITH CONTRACTORS DEBARRED, SUSPENDED, OR PROPOSED FOR DEBARMENT
52.212-8	DEFENSE PRIORITY AND ALLOCATION REQUIREMENTS
52.215-2	AUDIT AND RECORDS - NEGOTIATION
52.215-10	PRICE REDUCTION FOR DEFECTIVE COST OR PRICING DATA

52.215-12	SUBCONTRACTOR COST OR PRICING DATA
52.215-14	INTEGRITY OF UNIT PRICES
52.215-8	ORDER OF PRECEDENCE
52.215.18	REVERSION OR ADJUSTMENT OF PLANS FOR POSTRETIREMENT BENEFITS OTHER THAN PENSIONS
52.222-3	CONVICT LABOR
52.222-26	EQUAL OPPORTUNITY
52.222-35	AFFIRMATIVE ACTION FOR SPECIAL DISABLED AND VIETNAM ERA VETERANS
52.222-36	AFFIRMATIVE ACTION FOR HANDICAPPED WORKERS
52.222-37	EMPLOYMENT REPORTS ON SPECIAL DISABLED VETERAN AND VETERANS OF THE VIETNAM ERA
52.223-2	CLEAN AIR AND WATER
52.223-6	DRUG-FREE WORKPLACE
52.224-1	PRIVACY ACT NOTIFICATION
52.224-2	PRIVACY ACT
52.225-13	RESTRICTIONS ON CONTRACTING WITH SANCTIONED PERSONS
52.227-1	ALT. I - AUTHORIZATION AND CONSENT
52.227-2	NOTICE AND ASSISTANCE REGARDING PATIENT AND COPYRIGHT INFRINGEMENT

52.227-10	FILING OF PATENT APPLICATIONS - CLASSIFIED SUBJECT MATTER
52.227-11	PATENT RIGHTS - RETENTION BY THE CONTRACTOR (SHORT FORM)
52.228-7	INSURANCE - LIABILITY TO THIRD PERSONS
52.230-5	COST ACCOUNTING STANDARDS - EDUCATIONAL INSTRUCTIONS
52.232-23	ALT. I - ASSIGNMENT OF CLAIMS
52.233-1	DISPUTES
52.233-3	ALT. I - PROTEST AFTER AWARD
52.237-3	CONTINUITY OF SERVICES
52.246-25	LIMITATION OF LIABILITY - SERVICES
52.247-63	PREFERENCE FOR U.S. - FLAG AIR CARRIERS
52.249-5	TERMINATION FOR CONVENIENCE OF THE GOVERNMENT (EDUCATIONAL AND OTHER NONPROFIT INSTITUTIONS)
52.249-14	EXCUSABLE DELAYS
52.251-1	GOVERNMENT SUPPLY SOURCES

DOD FAR CLAUSES	DESCRIPTION
252.203-7001	SPECIAL PROHIBITION ON EMPLOYMENT
252.215-7000	PRICING ADJUSTMENTS
252.233-7004	DRUG FREE WORKPLACE (APPLIES TO SUBCONTRACTS WHERE THERE IS ACCESS TO CLASSIFIED INFORMATION)
252.225-7001	BUY AMERICAN ACT AND BALANCE OF PAYMENTS PROGRAM
252.225-7002	QUALIFYING COUNTRY SOURCES AS SUBCONTRACTS
252.227-7013	RIGHTS IN TECHNICAL DATA - NONCOMMERCIAL ITEMS
252.227-7030	TECHNICAL DATA - WITHOLDING PAYMENT
252.227-7037	VALIDATION OF RESTRICTIVE MARKINGS ON TECHNICAL DATA
252.231-7000	SUPPLEMENTAL COST PRINCIPLES
252.232-7006	REDUCTIONS OR SUSPENSION OF CONTRACT PAYMENTS UPON FINDING OF FRAUD

APPENDIX 2:

SAMPLE TECHNICAL EVALUATION FORM

SUMMER RESEARCH EXTENSION PROGRAMTECHNICAL EVALUATION

SREP NO: 95-0811
SREP PRINCIPAL INVESTIGATOR: Dr. Michael Burke

Circle the rating level number, 1 (low) through 5 (high), you feel best evaluate each statement and return thecompleted form by mail to:

RDL
Attn: 1995 SREP Tech Evals
5800 Uplander Way
Culver City, CA 90230-6608
(310) 216-5940 or (800) 677-1363

1.	This SREP report has a high level of technical merit.	1 2 3 4 5
2.	The SREP program is important to accomplishing the lab's mission.	1 2 3 4 5
3.	This SREP report accomplished what the associate's proposal promised.	1 2 3 4 5
4.	This SREP report addresses area(s) important to the USAF.	1 2 3 4 5
5.	The USAF should continue to pursue the research in this SREP report.	1 2 3 4 5
6.	The USAF should maintain research relationships with this SREP associate.	1 2 3 4 5
7.	The money spent on this SREP effort was well worth it.	1 2 3 4 5
8.	This SREP report is well organized and well written.	1 2 3 4 5
9.	I'll be eager to be a focal point for summer and SREP associates in the future.	1 2 3 4 5
10.	The one-year period for complete SREP research is about right.	1 2 3 4 5

---

11. If you could change any one thing about the SREP program, what would you change.

---

12. What would you definitely NOT change about the SREP program?

USE THE BACK FOR ANY ADDITIONAL COMMENTS.

Laboratory: Armstrong Laboratory
Lab Focal Point: Linda Sawin Office Symbol: AL/HRMI
Phone: (210) 536-3876

James Anderson report unavailable at time of publication.

FINITE ELEMENT MODELING OF THE HUMAN NECK AND ITS VALIDATION FOR THE ATB

Dr. Hashem Ashrafiuon
Associate Professor
Department of Mechanical Engineering

Villanova University
Villanova, PA 19085

Final Report for:
Summer Research Extension Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

and

Villanova University

December 1995

FINITE ELEMENT MODELING OF THE HUMAN NECK AND ITS VALIDATION FOR THE ATB

Hashem Ashrafiuon
Associate Professor
Department of Mechanical Engineering
Villanova University

Abstract

The Articulated Total Body (ATB) is a rigid body dynamic model of the human body used at the Armstrong Laboratory. The model is used to determine the mechanical response of the human body in different dynamic environments such as aircraft pilot ejections, sled tests, etc. The new version of the ATB model program is capable of treating the individual segments of the human body or other systems as deformable segments. The model assumes that the natural frequencies, damping characteristics and vibration normal modes of the deformable segments have been determined in advance. A particular segment of the human body which may undergo significant deformation is the human neck. Finite element models of several human necks are developed. Solid finite elements with linearly elastic isotropic material properties are used for the models. The finite element analysis results (frequencies and mode shapes) are then incorporated into the ATB model for simulations. The simulation results are compared with several test results available at the Armstrong Laboratory to verify the validity and accuracy of the models.

FINITE ELEMENT MODELING OF THE HUMAN NECK AND ITS VALIDATION FOR THE ATB

Hashem Ashrafiuon

Introduction

The Articulated Total Body (ATB) model is used at the Armstrong Laboratory (AL) for predicting gross motion of the human body under various dynamic environments. The new modifications to the ATB model allows the individual segments of the human body and other parts involved in the simulation to be modeled as rigid and/or deformable segments (Ashrafiuon, 1993).

It is assumed that the displacement of a deformable segment relative to a reference frame attached to its undeformed state is small and linear. Therefore, linear elastic displacement field of a segment may be defined by linear combinations of its vibration normal modes. In order reduce the size and complexity of the problem, a small number of modes may be selected to approximate the displacement field using Ritz approximation (Wilson, et al. 1982). The vibration normal modes along with natural frequencies and damping characteristics of each deformable segment are determined using finite element modal analysis.

In the case of human body, the segment that is expected to undergo significant deformation is the human. Finite element models of the neck for several human subjects are developed. The natural frequencies, modal damping, vibration modes of the deformable are determined and incorporated into the new ATB model. The models are then validated through ATB model simulations and comparison with data obtained from the sled tests available at Armstrong Laboratory. The process is similar to the one explained in detail for Hybrid III (Kaleps, et al. 1988) manikin necks by Ashrafiuon (1994) and Colbert (1994). The mathematical formulation has been presented in detail by Ashrafiuon (1993).

Human Neck Model

The Naval Biodynamics Development Laboratory (NBDL) has provided anthropometric data as well as sled test data for eight subjects. Of these eight subjects, four were chosen to model with the ATB using the deformable body capability. The subjects selected are numbered 127, 130, 133, 134. These four were chosen primarily because there was complete data sets available. An ATB model for the NBDL sled tests was previously developed by AL/CFBV. These simulations were developed using a three body model consisting of the head, neck, and the upper torso. This model uses only rigid bodies. Comparisons of the NBDL sled test data will be made using the original ATB model, an improved rigid ATB model, and an ATB model using a deformable neck for each subject at four different acceleration levels (4g, 8g, 10g, 12g).

For each of the four subjects chosen for this simulation, the neck characteristics must be modeled with the ATB and with finite elements. Comparisons of the original ATB input files provided by AL/CFBV with the NBDL anthropometric data show that the original ATB model uses much shorter neck lengths than the actual lengths provided

by NBDL. In addition, several other inconsistencies exist in the ATB input file. For example, initial torques are provided to the head and neck pins and the joint characteristics seem inconsistent with physical reality.

Before any deformable capabilities were added, an improved ATB rigid model was developed. The neck length was not explicitly listed in the NBDL anthropometric data, however an approximation was made for neck length using the measured characteristic labeled 'T1 to RT Tragion'. The Tragion height is very close to the occipital condyle joint height which is represented by the head pin joint in the ATB. Other properties such as segment weights and inertias were obtained from the original ATB input file. With the neck lengths determined for each subject, the ATB joint data was modified such that the original height of the head center of gravity (cg) to the neck pin was maintained. In the original ATB file, the neck length was quite short for each subject while the distance from the head cg to the head pin was rather long. Using the corrected neck lengths while maintaining the same overall height effectively shortens the head segment. Table 1 summarizes the neck lengths as obtained from the NBDL data and the neck weights from the original ATB input files.

Table 1. Human neck data summary

Subject	Neck Length (in)	Neck Weight (lb)
127	4.68	2.57
130	4.76	2.85
133	4.09	2.46
134	3.27	2.89

The initial conditions for each subject at the four different sled tests were determined from the experimental data provided by NBDL. Other experimental data that was provided by NBDL include head rotations and accelerations as well as the acceleration experienced by the sled and torso. For this simple three body model, the upper torso acceleration from the NBDL experimental results was used as the input in the ATB model simulations. Sled acceleration data is not applicable in these three body models.

The joint characteristics for the head and neck pin that needed to be adjusted in the original ATB model are those parameters that represent joint stiffness, damping, and energy loss or hystersis effects. It was the determination of these parameters which proved to be the greatest challenge in this improved ATB model. Three terms are needed to define the cubic polynomial which represents spring's elasticity while only one parameter defines the linear damping characteristics of the joint. Energy dissipation also has one parameter to modify for each joint.

A parametric study was performed to determine the correct combination of these characteristics to better correlate the ATB model simulations and the experimental results. Intuitively, larger stiffness and damping values were selected for the neck pin than the head pin. Since there are four subjects and four acceleration levels, the parametric study was

limited to tuning the parameters for good agreement at one acceleration level; namely the 10g acceleration level. Simulation results were plotted against the NBDL sled test data for the head rotation and the head x-acceleration. Parameters were varied until reasonably good agreement was obtained for all four subjects at the 10g level. Then, the same parameters were used for the other acceleration levels.

After completing the parametric study for the rigid ATB model, an ATB model was developed with a deformable neck. The deformable ATB model used the same joint parameters as determined from the parametric study. The finite element method was used to develop the neck models to determine the mode shapes and natural frequencies of the human necks which are later incorporated into the ATB model. The finite element neck models were deliberately kept simple to avoid unknown effects and also to limit the number of new parameters which may need variations. An example of the finite element model of the human neck is shown in Figure 1. The model has a cylindrical shape with inserts at the top and bottom of the cylinder to facilitate joint connections in the ATB model. Brick element are used throughout except at the joint attachment inserts which are meshed with tetrahedral elements. The nodes at the base (bottom) of the neck are fixed. Four finite element models were created, one for each subject. The models were developed using the length as provided by the NBDL data and the diameter was adjusted to match the weights of the necks as provided by the original ATB models. The material used for the neck is somewhat arbitrary. The density was varied in order to match the model weight while maintaining reasonable diameters for the neck, approximately two inches. The Young's Modulus of the material and the modal damping ratio were also varied in a parametric study in order to provide the best agreement with the experimental data. A Young's Modulus of 2000 psi and a modal damping ratio of 0.3 were selected which were used for all four finite element models. Only the first two mode shapes were used in the ATB sled test model and they are similar to the first two bending modes of a cantilevered beam. Table 2 provides a summary of the finite element models of the human necks.

Table 2. Finite element model summary

Subject	Nodes	Elements	Natural Frequency
127	1302	892	77.6 Hz
130	1302	892	77.4 Hz
133	1198	810	99.4 Hz
134	1674	1184	151.2 Hz

Results and Discussion

All sled test results are summarized in figs. 2-33. These figures include experimental data from four subjects at four different acceleration levels. For each simulation, experimental results of head rotation and head x-acceleration are compared with the ATB deformable model. For the 10g tests (figs. 6-7, 14-15, 22-23, 30-31), experimental results are

compared with the original rigid model, the improved rigid model, and the deformable model. This complete comparison is only presented for the 10g case since parametric study was done only at this acceleration level. It can be seen from these figures that the deformable model is a much more accurate representation of the system than any rigid model.

The parameters that provided the best agreement with the NBDL experimental data obtained from subject 130's 10g tests are summarized in Table 3. Three stiffness values are listed for the linear, quadratic, and cubic terms, respectively. Examination of the 10g acceleration level plots (figs. 6-7, 14-15, 22-23, 30-31) show good to excellent agreement with the experimental results and obviously a drastic improvement over the original ATB model of the NBDL sled tests for all subjects. This should be expected since the parametric study was tuned at this acceleration level. Since one of the objectives of this program was to provide a uniform set of parameters to be used for most subjects, one set of parameters was developed and then applied to the other subjects at the other acceleration levels. The rest of the plots show average to good agreement with the experimental data and in a few instances, the results exhibit poor to fair agreement with the NBDL sled tests.

Variations in different subjects is to be expected especially when only one set of parameters is used for all subjects. However, the results suggest that there are variations for each subject at the various acceleration levels. Possible explanations may be erroneous test results or the variation in the repeatability of the tests. Another explanation for the variations could be that at different impact speeds, a particular subject simply may react differently by tensing muscles or adjusting position. This explanation is hard to quantify. Probably the best explanation for the variations in the results for the different subjects at the different accelerations is simply diversity. The bell curve governs height, weight, intelligence, and many other factors. Therefore it is only possible to state an average height for humans as well as an average stiffness parameter for the neck joint. Parameter variations could be made for each subject and acceleration level to provide better agreement but this is not practical.

Table 3. Summary of joint parametric study

Joint	Stiffness Parameters	Damping Parameter	Energy Dissipation
NP	75, 0.0, 0.1	1.3	0.5
HP	0.5, 0.0, 0.015	0.25	0.3

Conclusions

Finite element model of several human subjects have been developed and their natural frequencies and mode shapes are determined. The finite element results were successfully incorporated into the new deformable ATB model and several simulations at different acceleration levels were performed for each subject. The results clearly show two conclusions; the deformable ATB model predicts dynamic behavior better than a purely rigid body model, and that choosing realistic parameters for a model will always exhibit better results.

References

Ashrafiuon, H., 1994, "ATB Simulation of Deformable Manikin Neck Models," The 1994 AFOSR Summer Faculty Research Program, Armstrong Laboratory, Final Report.

Ashrafiuon, H., 1993, "Modeling of Flexible Bodies for the ATB Model," The 1993 AFOSR Summer Faculty Research Program, Armstrong Laboratory, Final Report.

Colbert, R., 1994, "Finite Element Modeling of Manikin Necks for the ATB Model," The 1994 AFOSR Graduate Student Research Program, Armstrong Laboratory, Final Report.

Kaleps, I., White, R.P., Beecher, R.M., Whitestone, J., and Obergefell, L.A., 1988, "Measurement of Hybrid III Dummy Properties and Analytical Simulation Data Base Development," Report No. AAMRL-TR-88-05, National Technical Information Service, Springfield, Virginia.

Wilson, E. L., Yuan, M-W, and Dickens, J. M., 1982, "Dynamic Analysis by Direct Superposition of Ritz Vectors," Earthquake Engineering and Structural Dynamics, Vol. 10, pp. 813-821.

Figure 1. Finite element model of a human neck

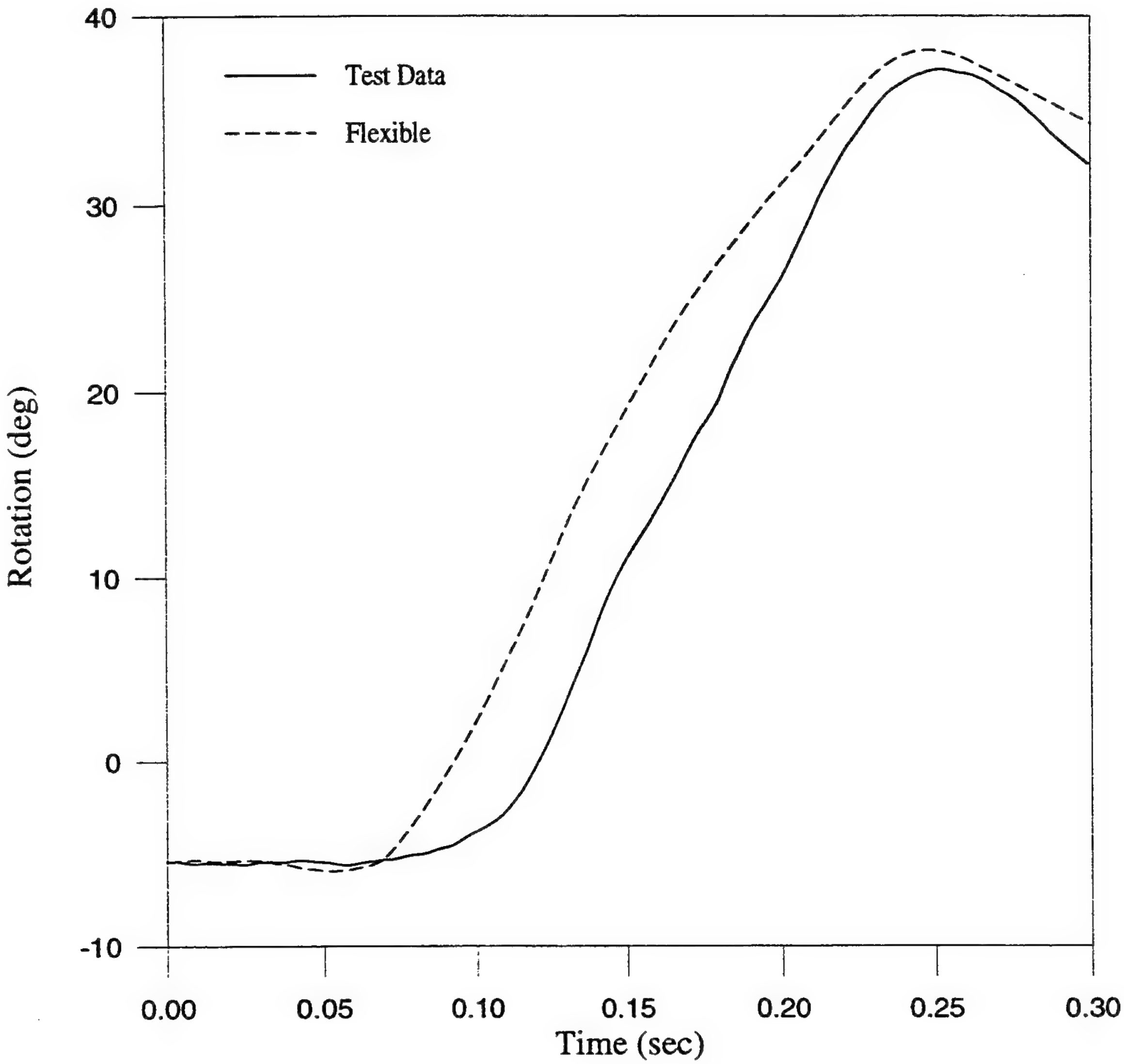


Figure 2. Subject 127 head rotation in the 4g sled test

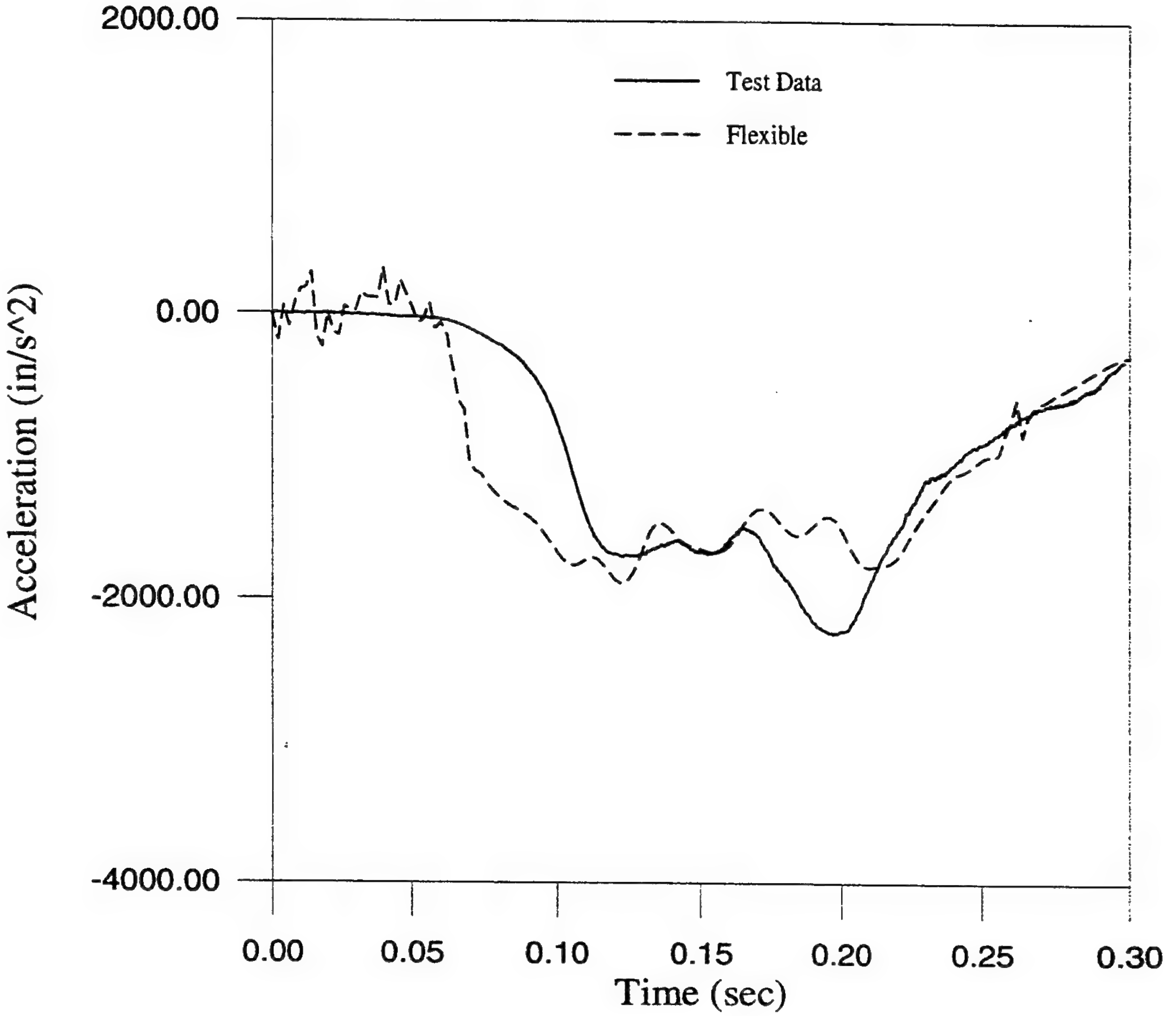


Figure 3. Subject 127 head x-acceleration in the 4g sled test

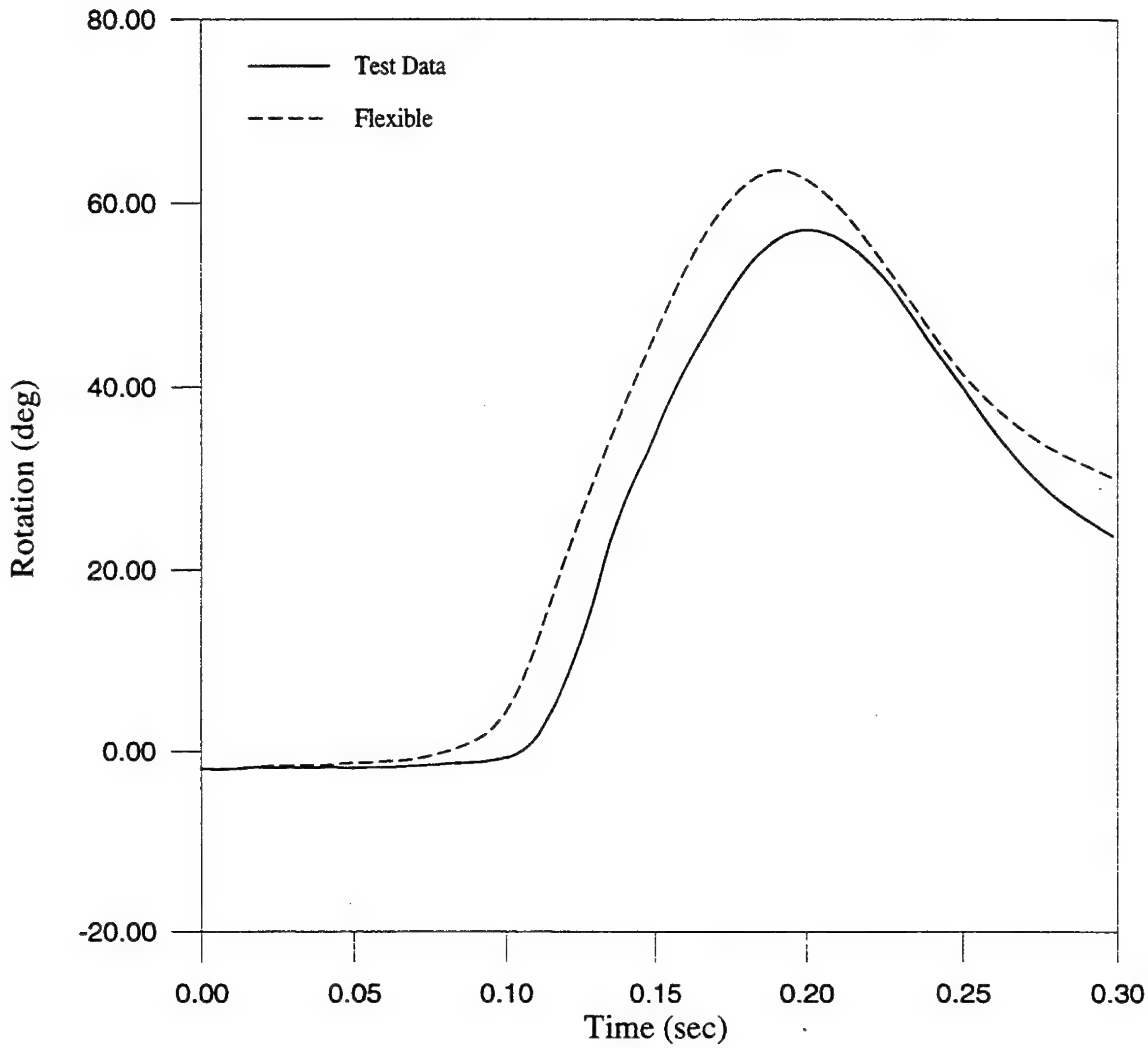


Figure 4. Subject 127 head rotation in the 8g sled test

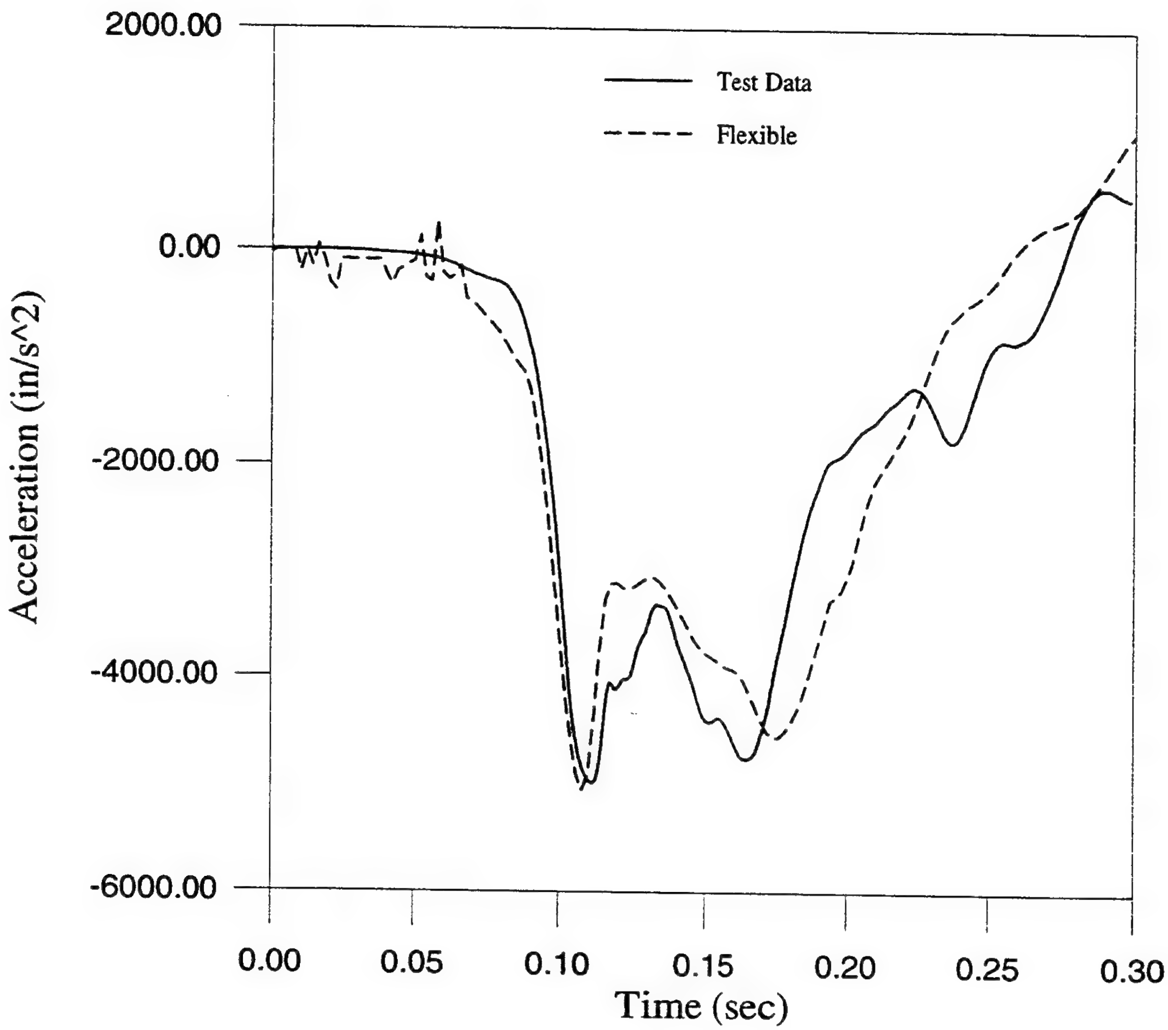


Figure 5. Subject 127 head x-acceleration in the 8g sled test

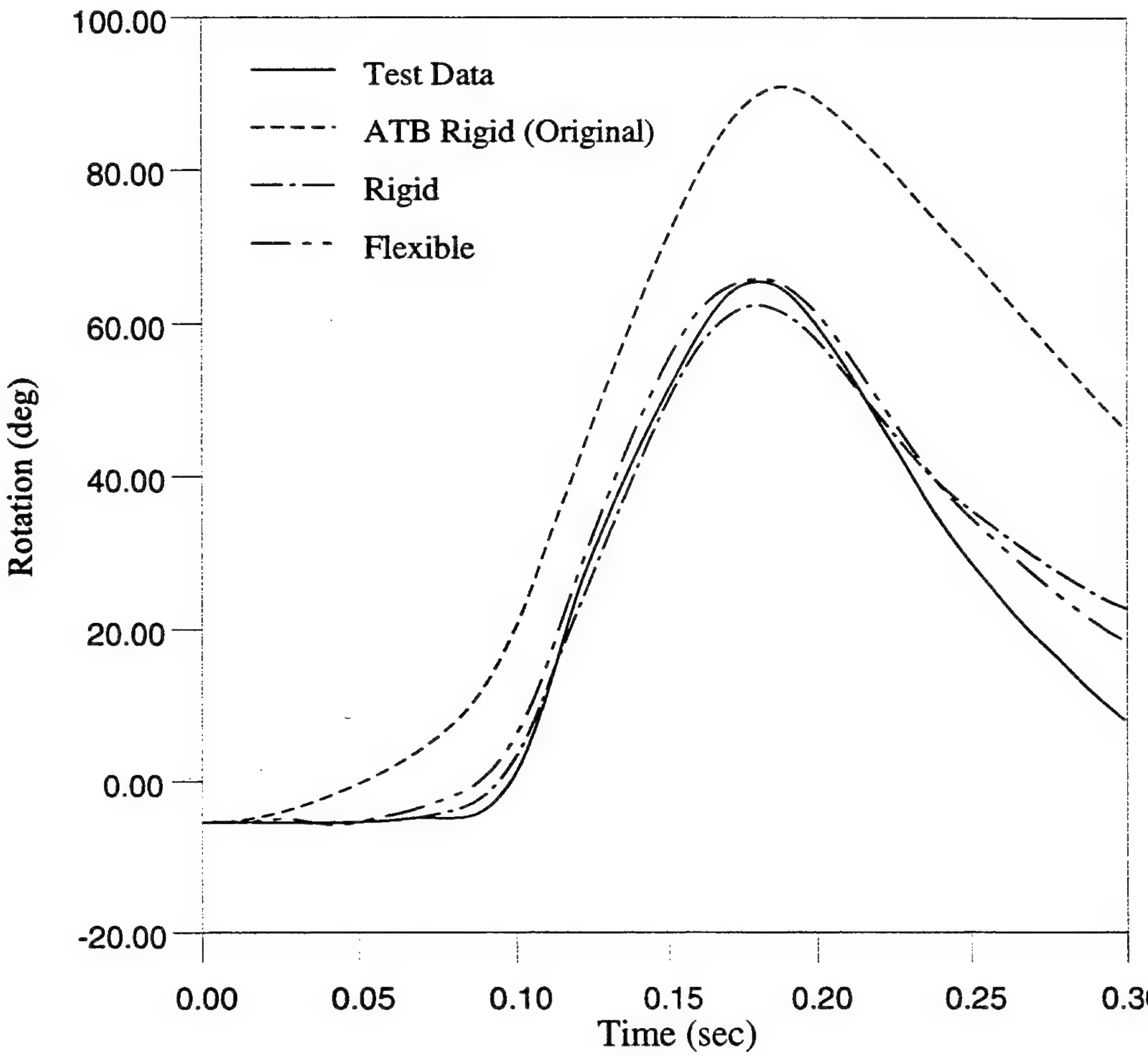


Figure 6. Subject 127 head rotation in the 10g sled test

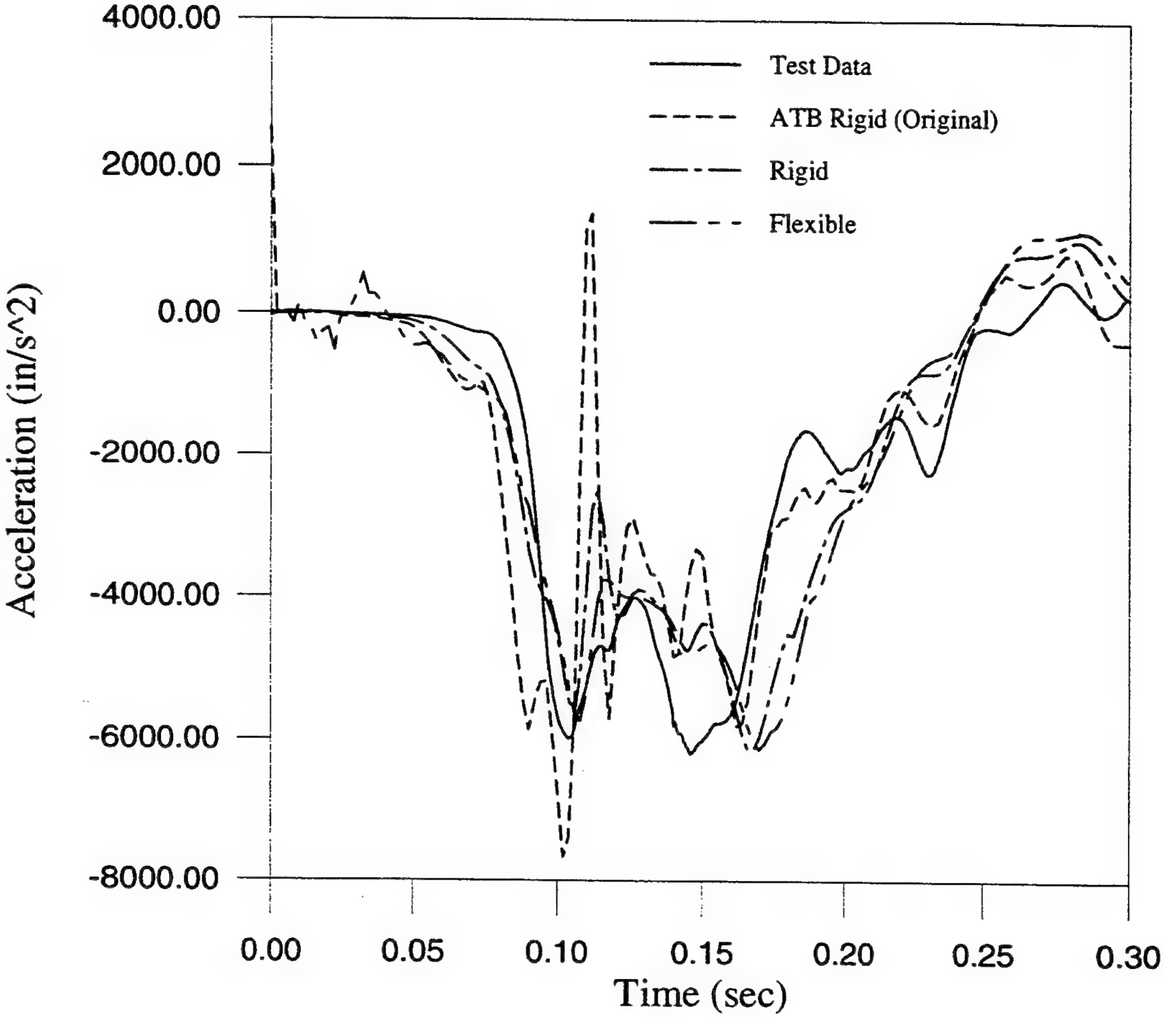


Figure 7. Subject 127 head x-acceleration in the 10g sled test

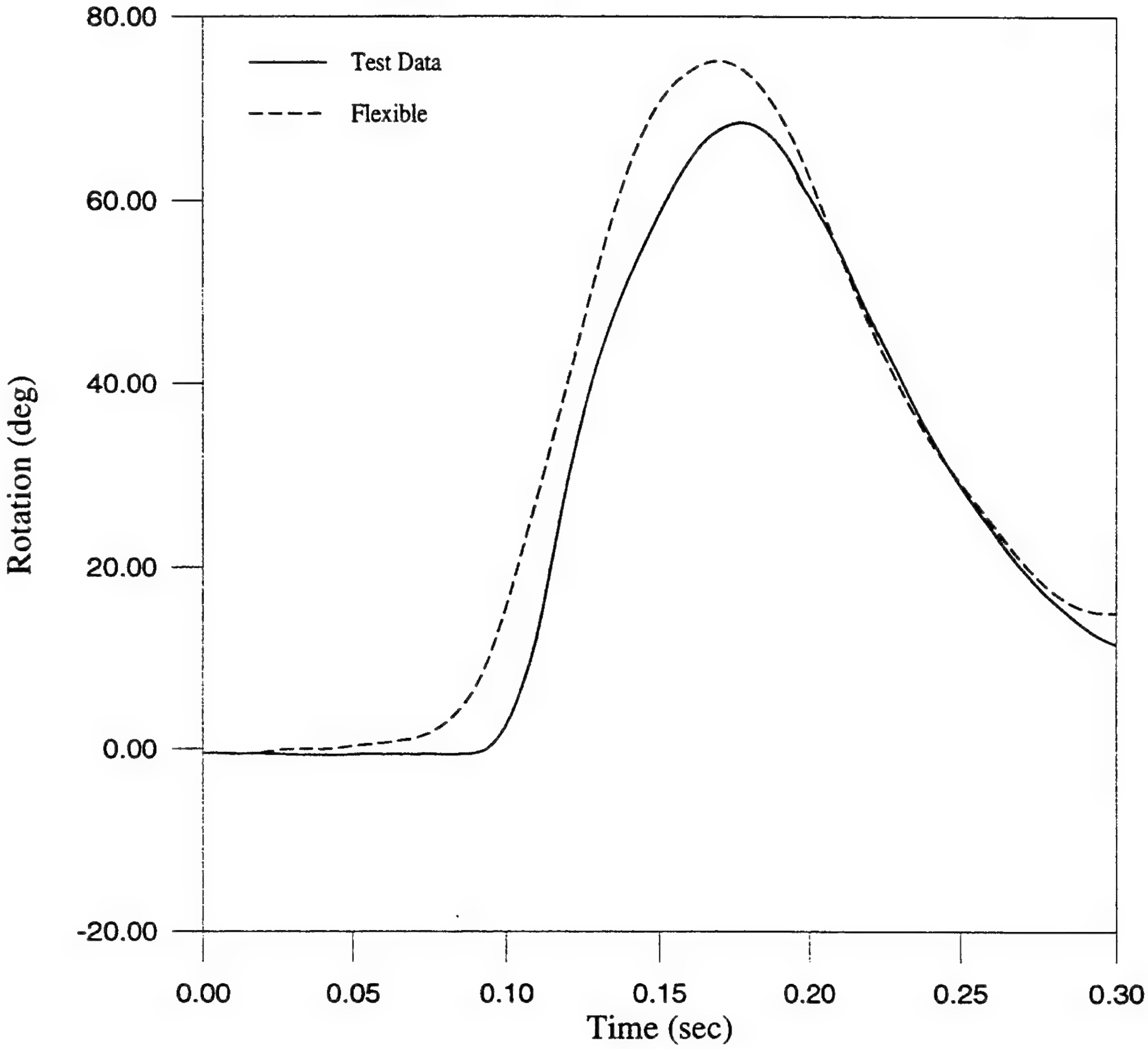


Figure 8. Subject 127 head rotation in the 12g sled test

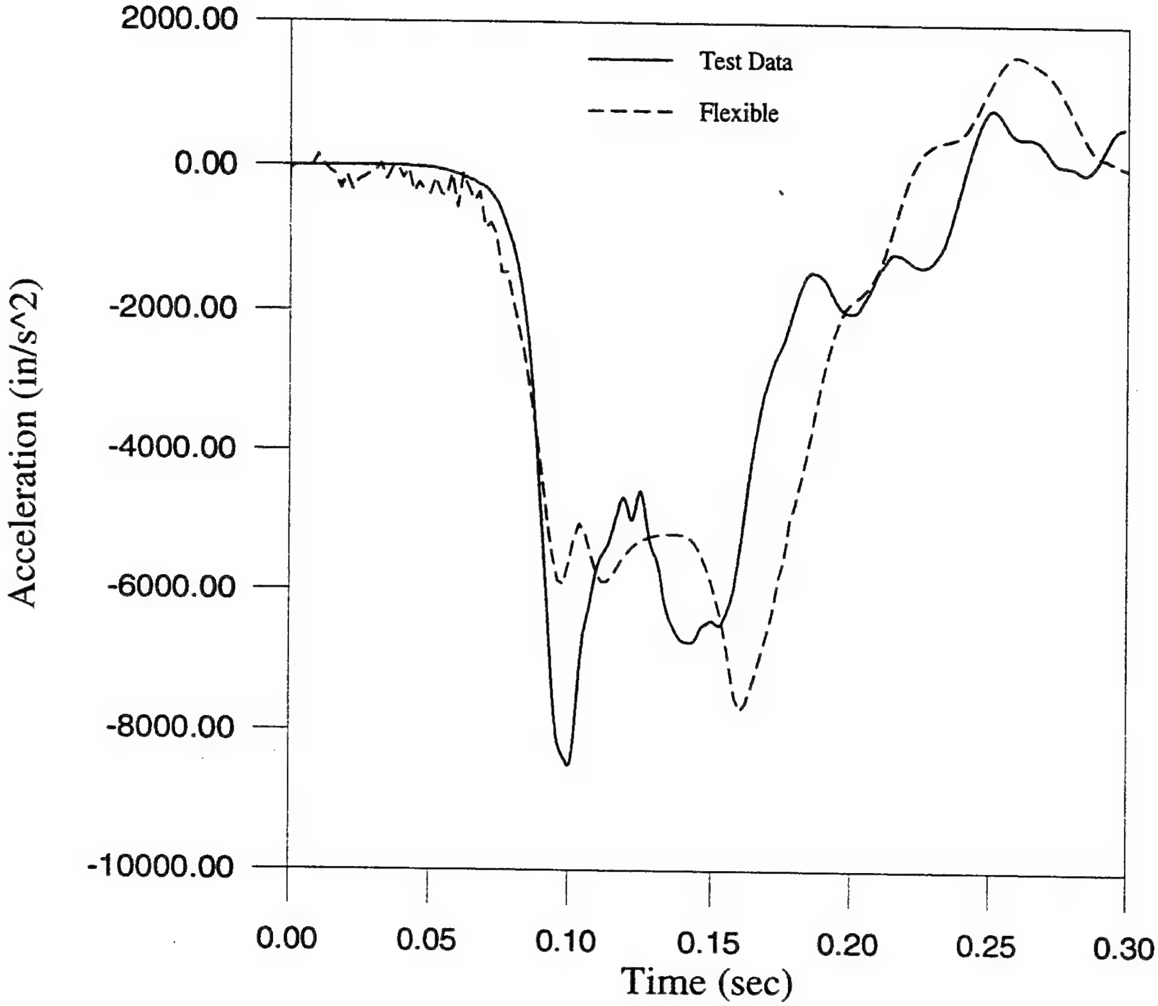


Figure 9. Subject 127 head x-acceleration in the 12g sled test

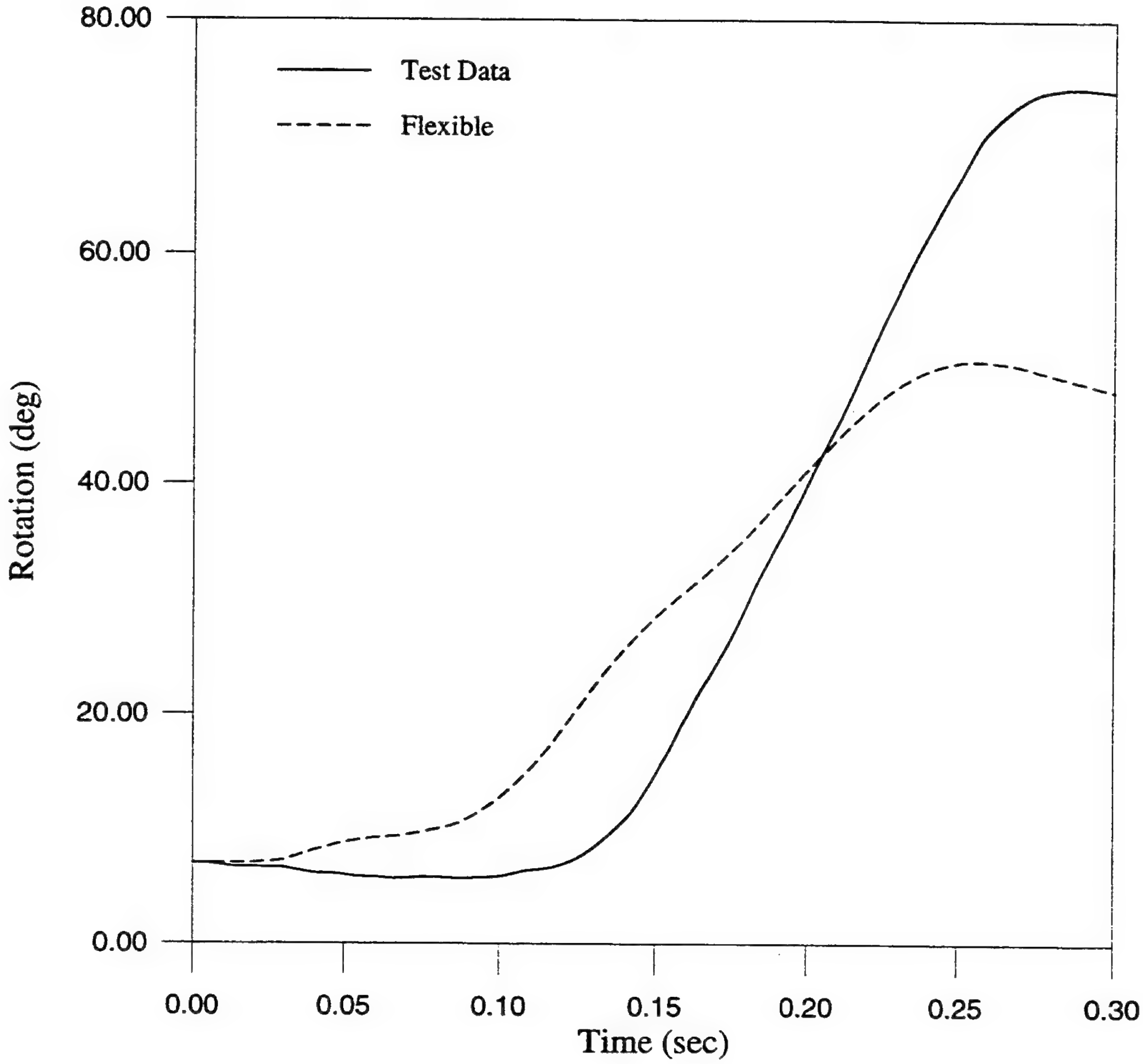


Figure 10. Subject 130 head rotation in the 4g sled test

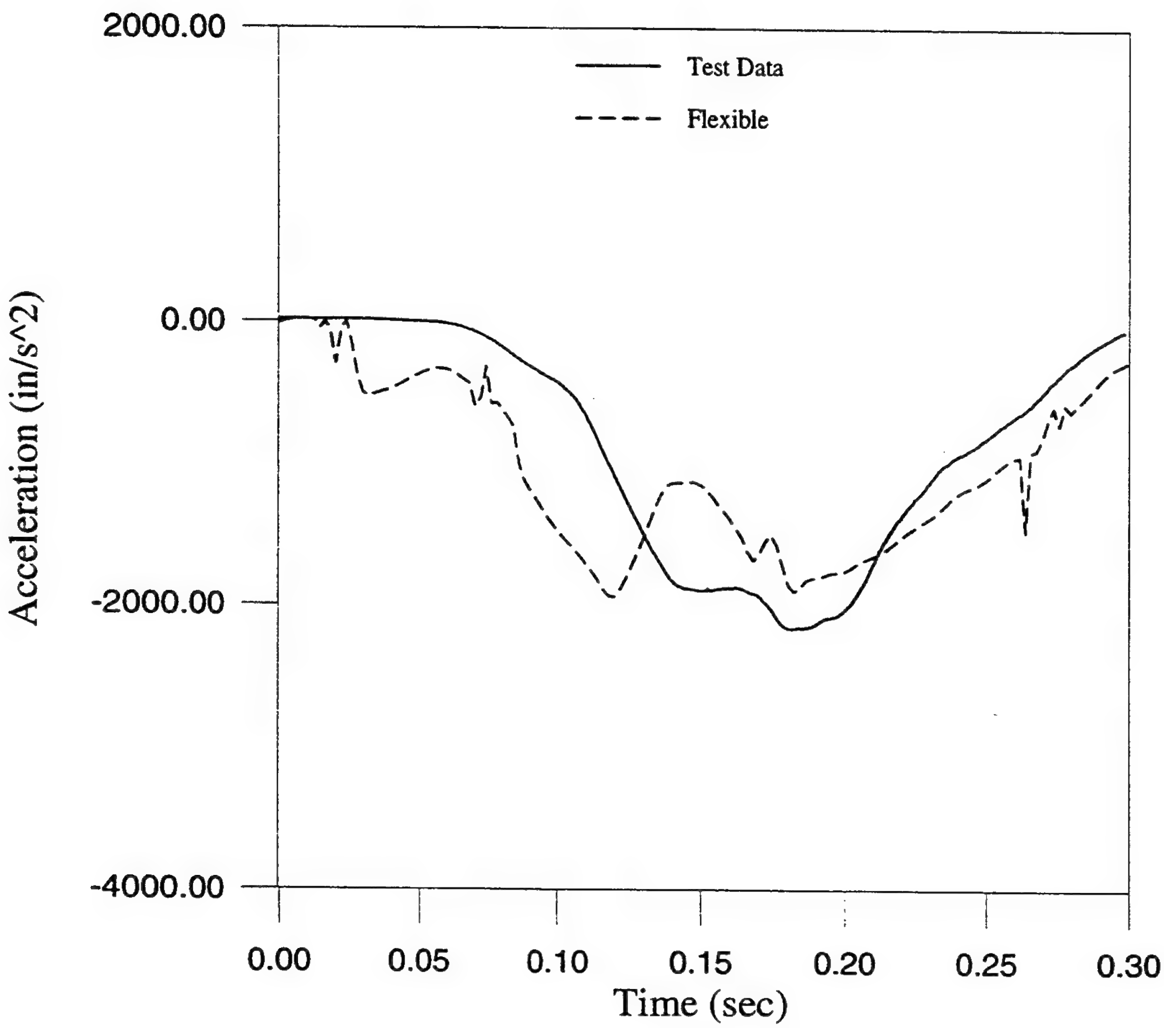


Figure 11. Subject 130 head x-acceleration in the 4g sled test

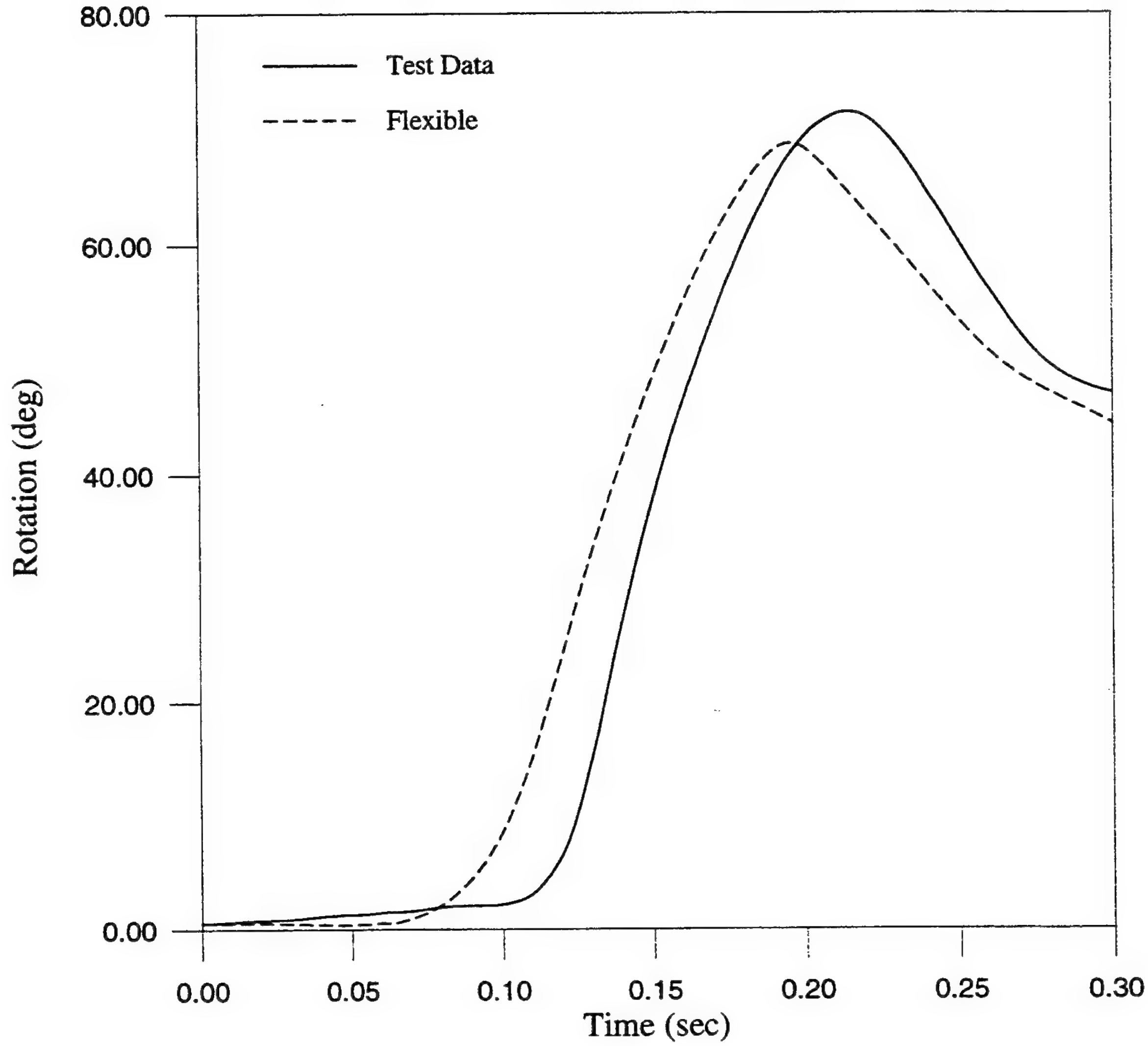


Figure 12. Subject 130 head rotation in the 8g sled test

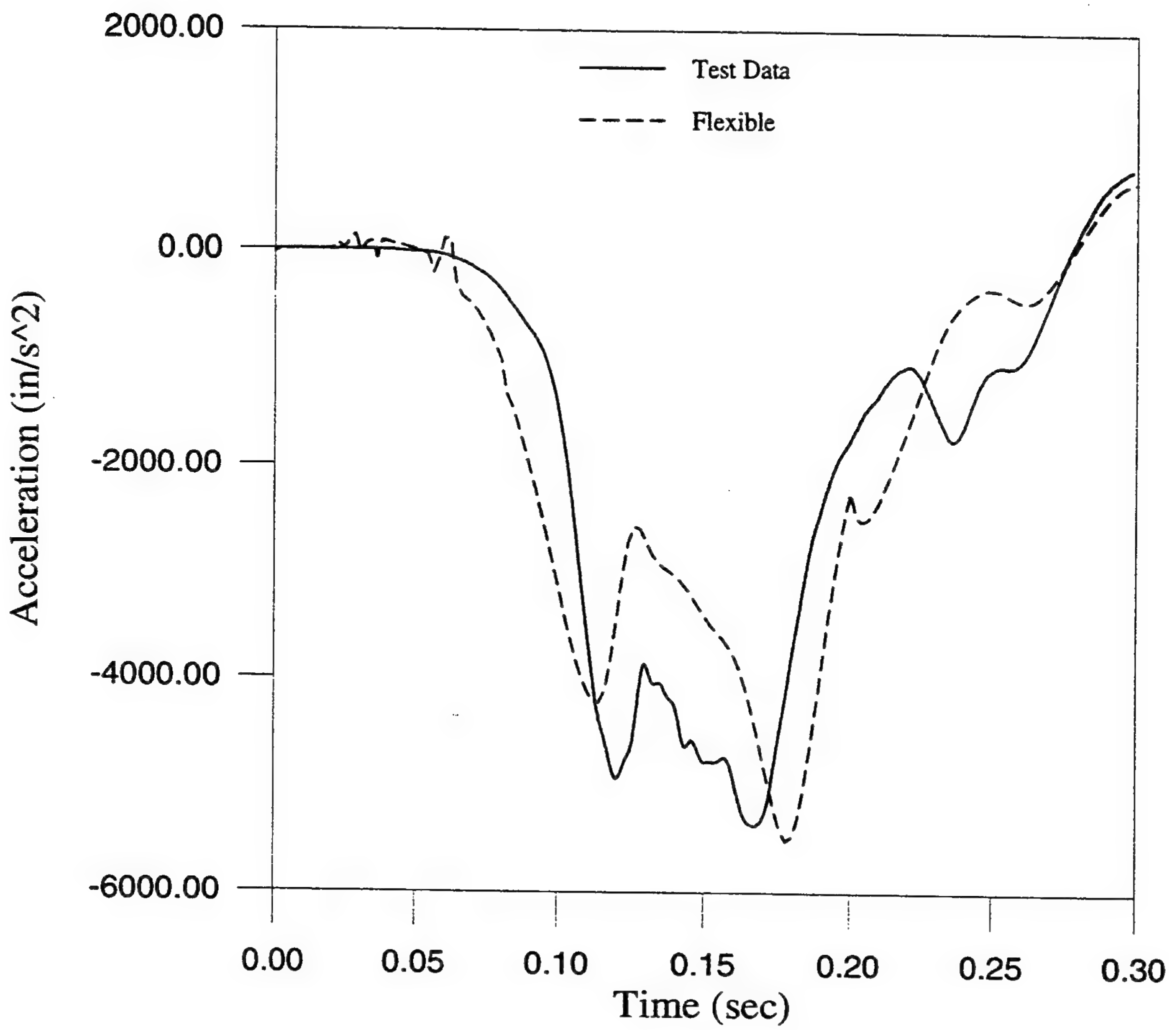


Figure 13. Subject 130 head x-acceleration in the 8g sled test

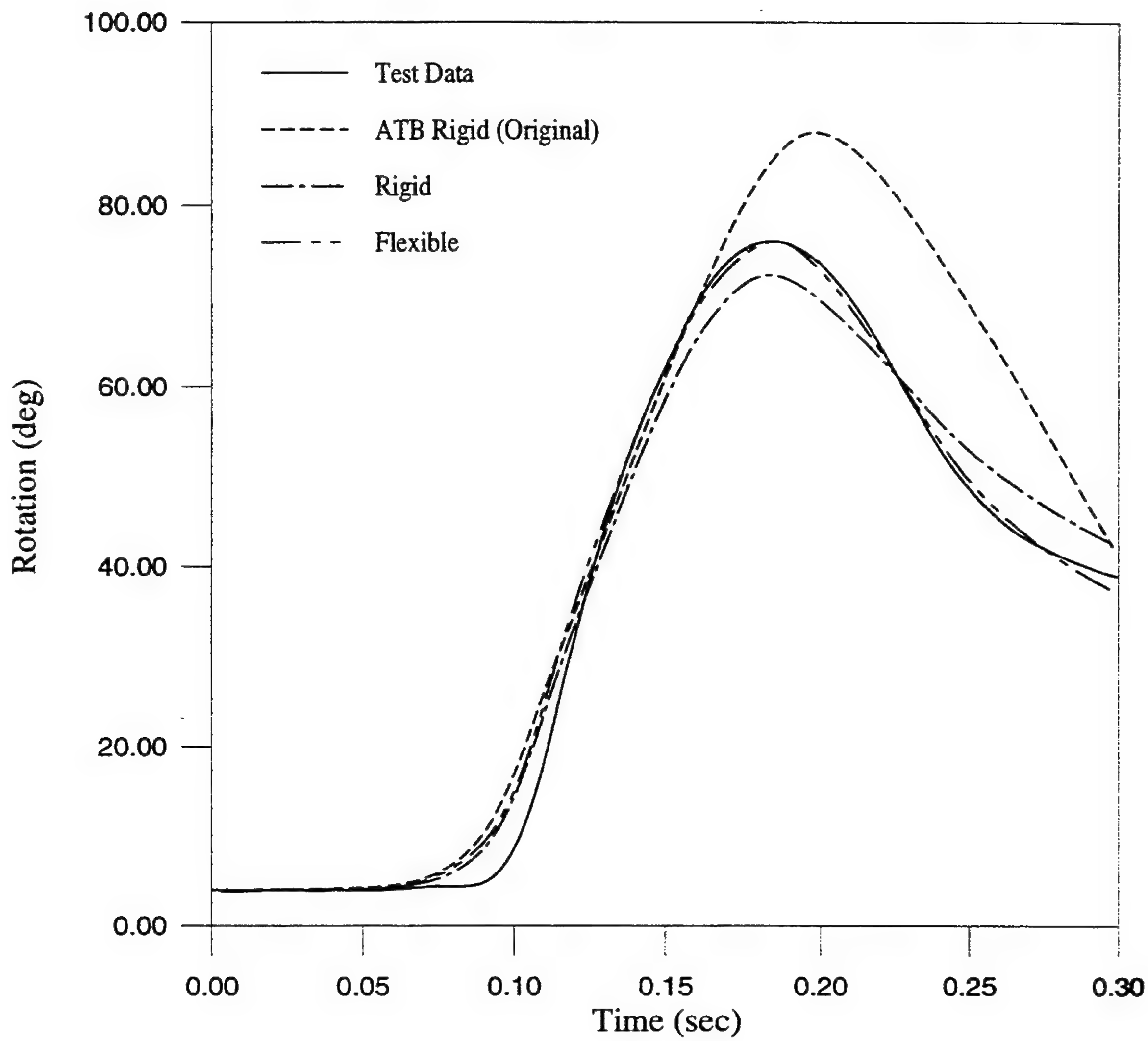


Figure 14. Subject 130 head rotation in the 10g sled test

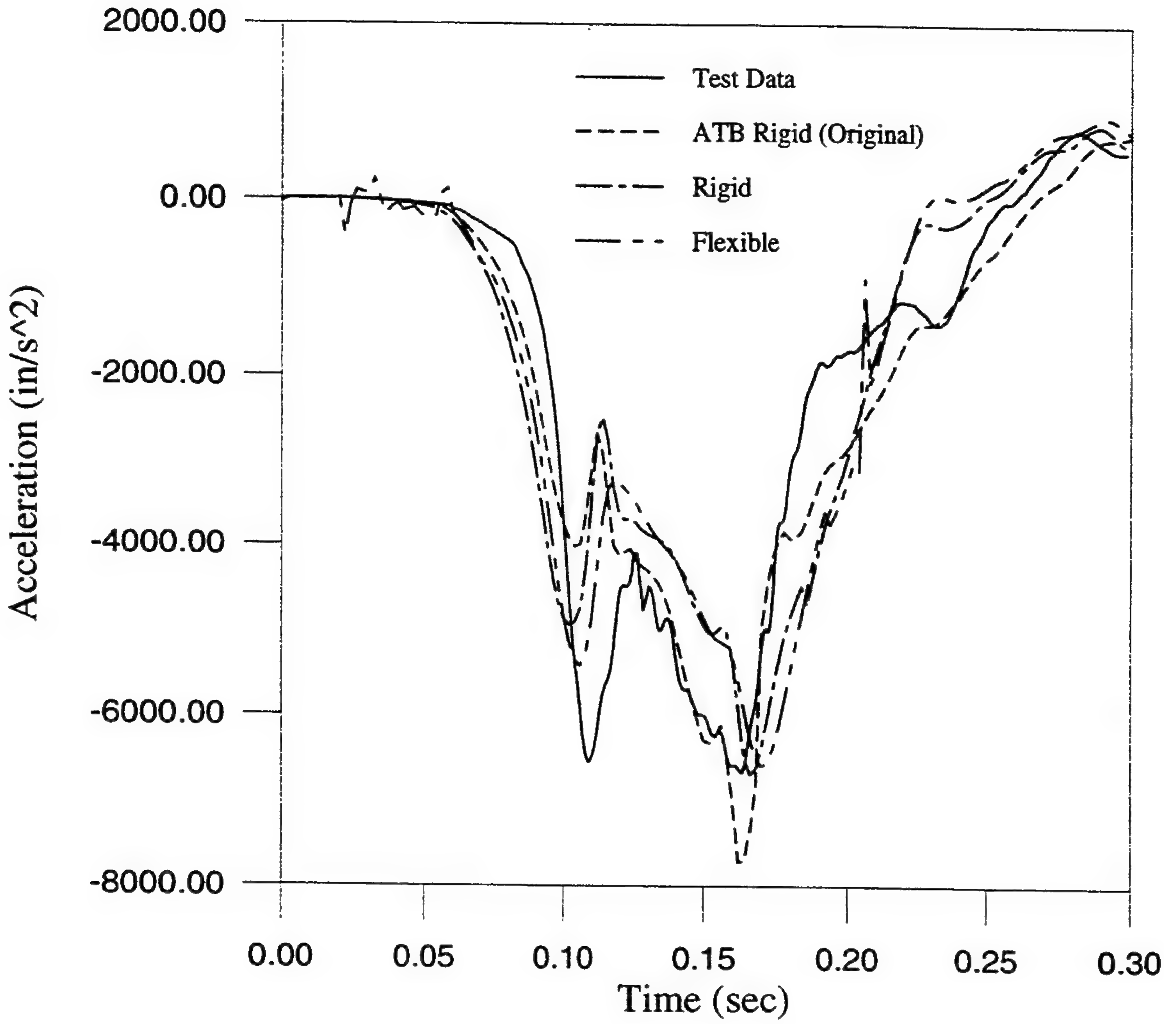


Figure 15. Subject 130 head x-acceleration in the 10g sled test

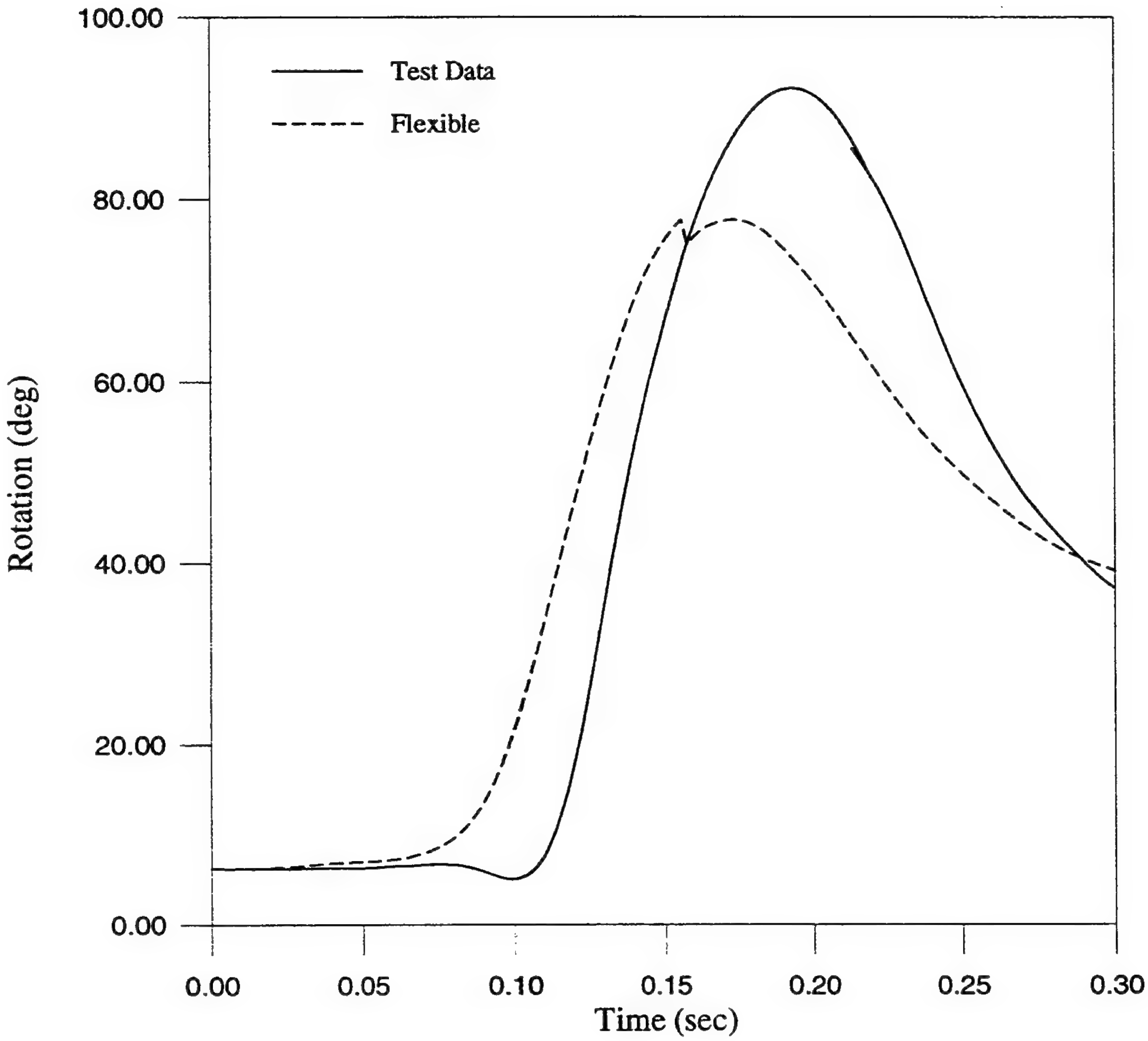


Figure 16. Subject 130 head rotation in the 12g sled test

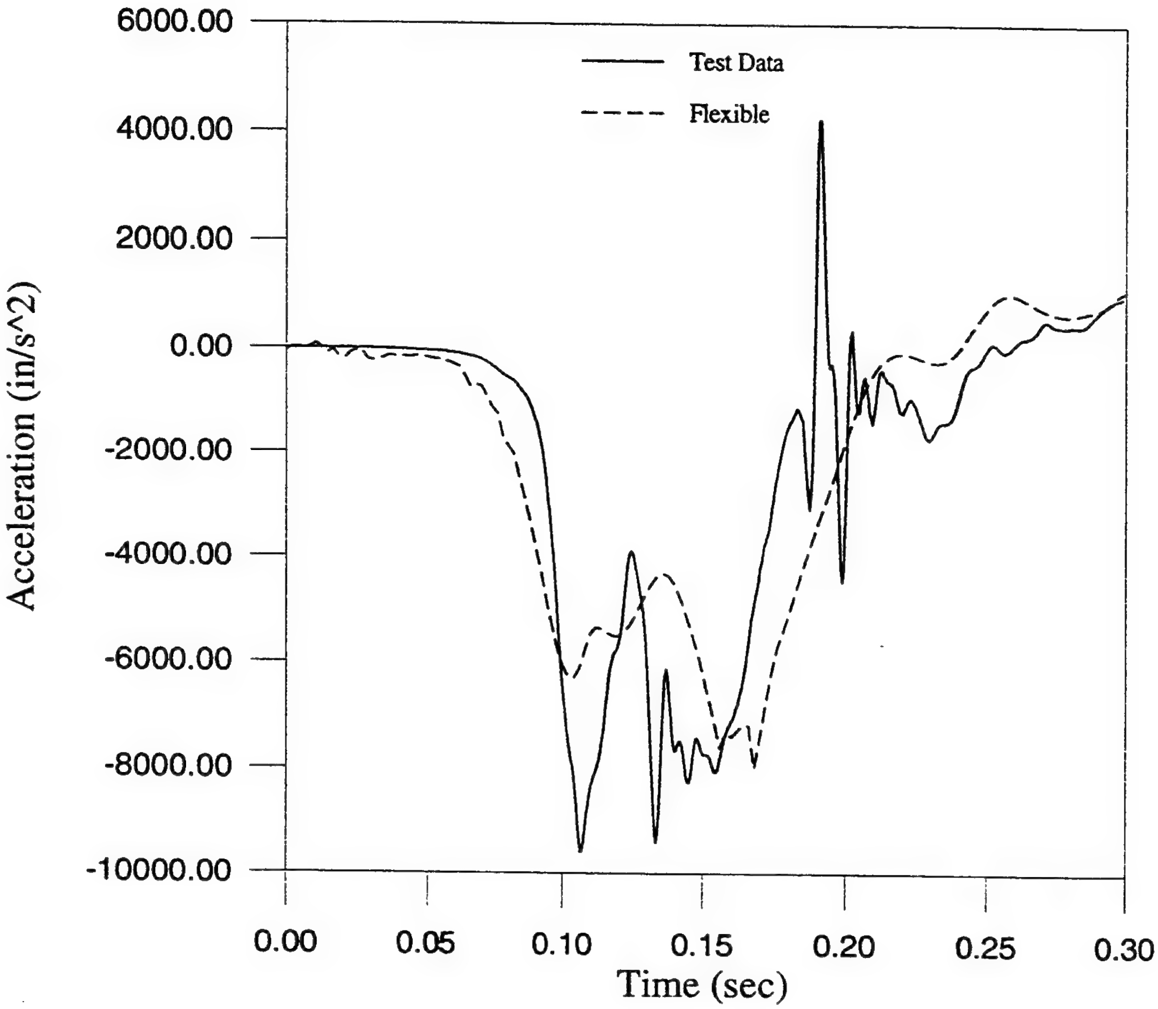


Figure 17. Subject 130 head x-acceleration in the 12g sled test

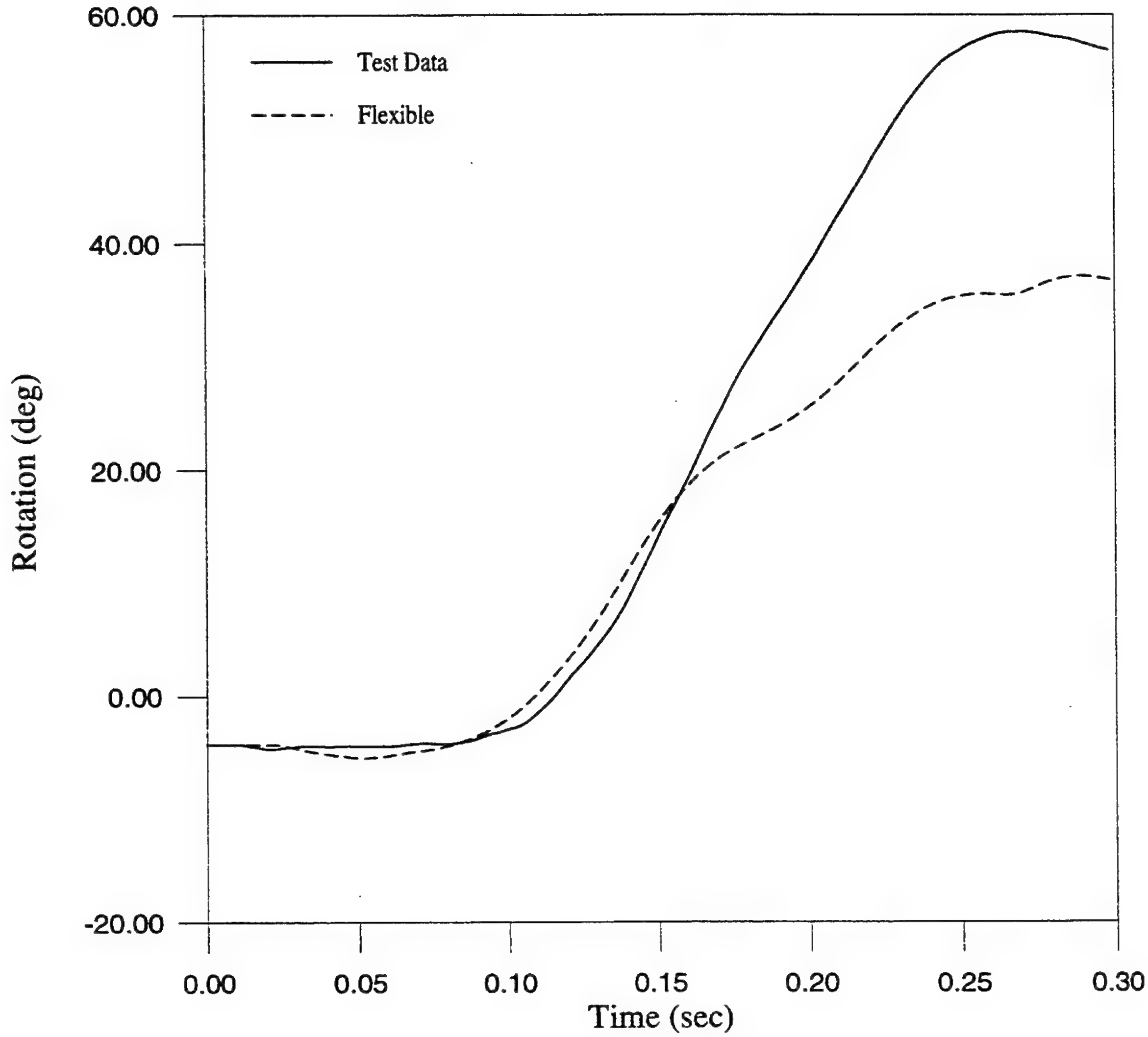


Figure 18. Subject 133 head rotation in the 4g sled test

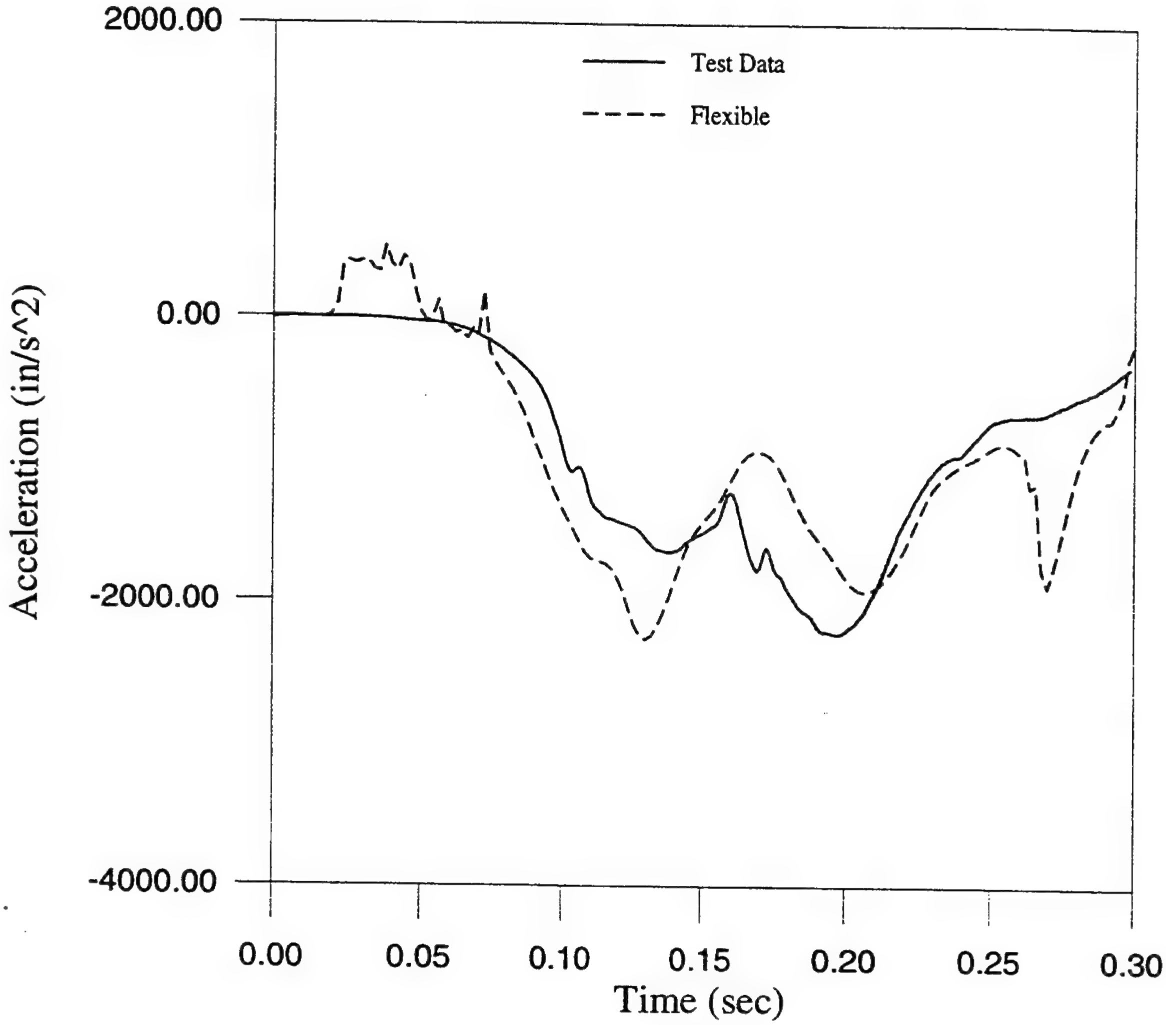


Figure 19. Subject 133 head x-acceleration in the 4g sled test

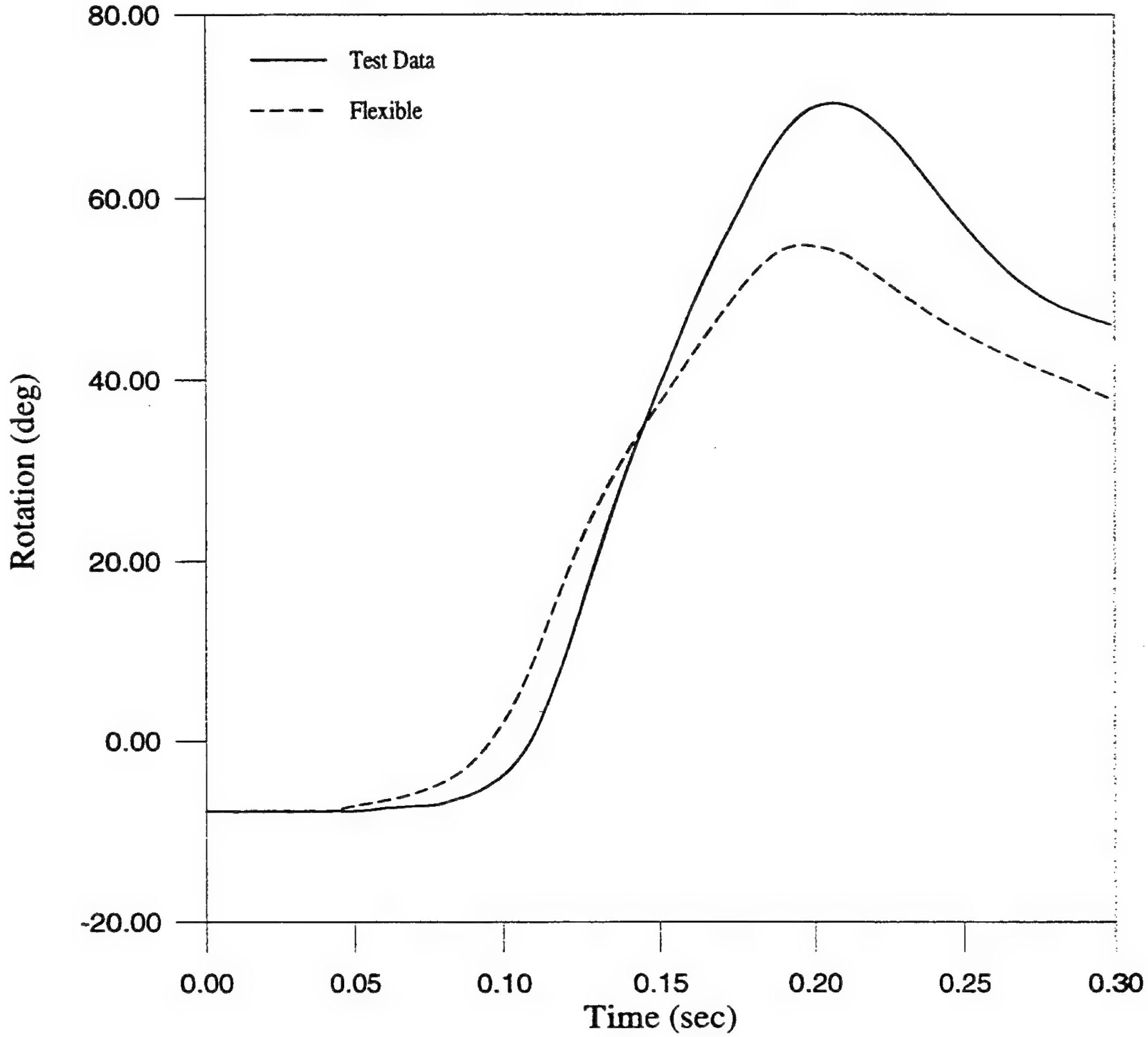


Figure 20. Subject 133 head rotation in the 8g sled test

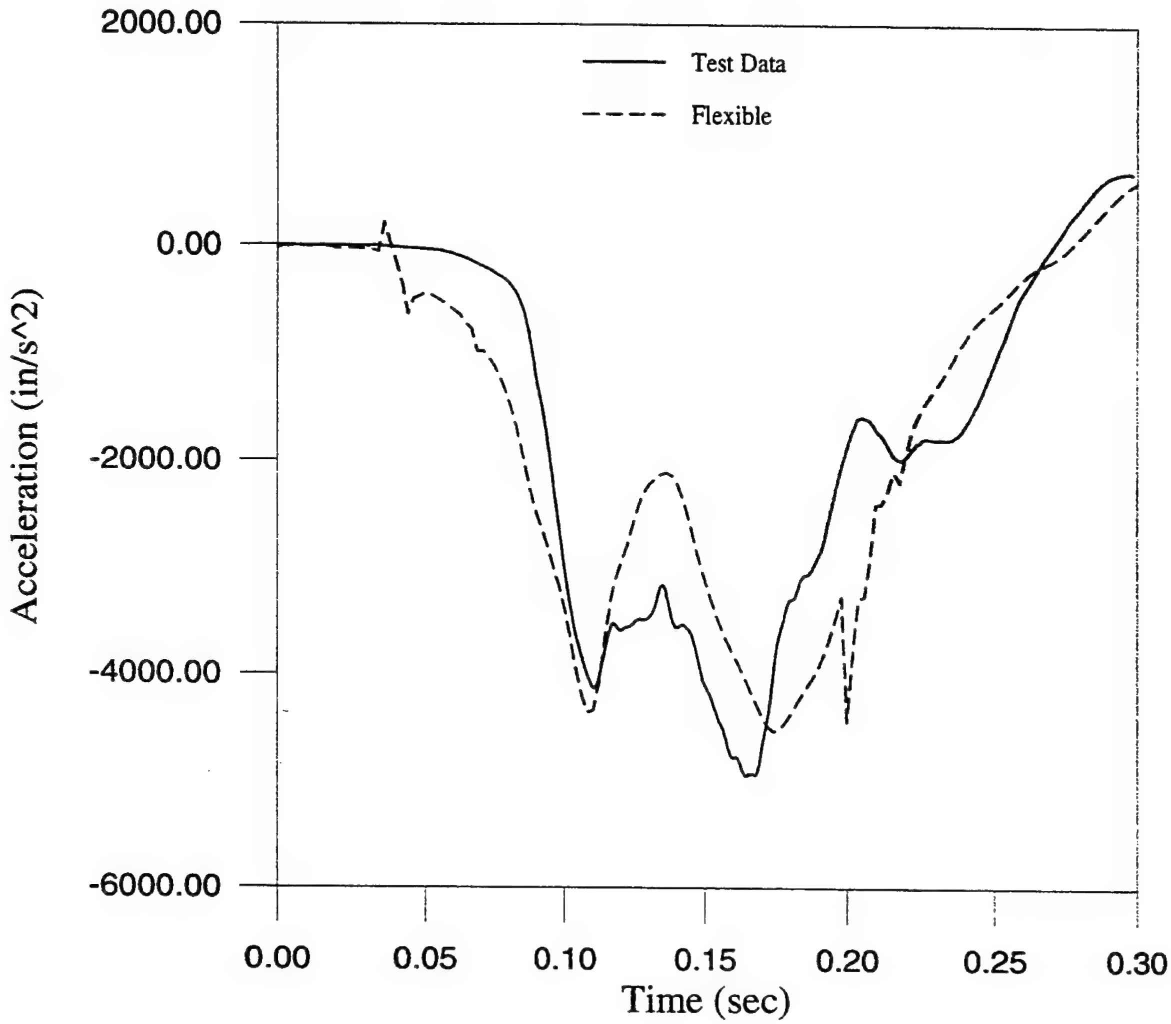


Figure 21. Subject 133 head x-acceleration in the 8g sled test

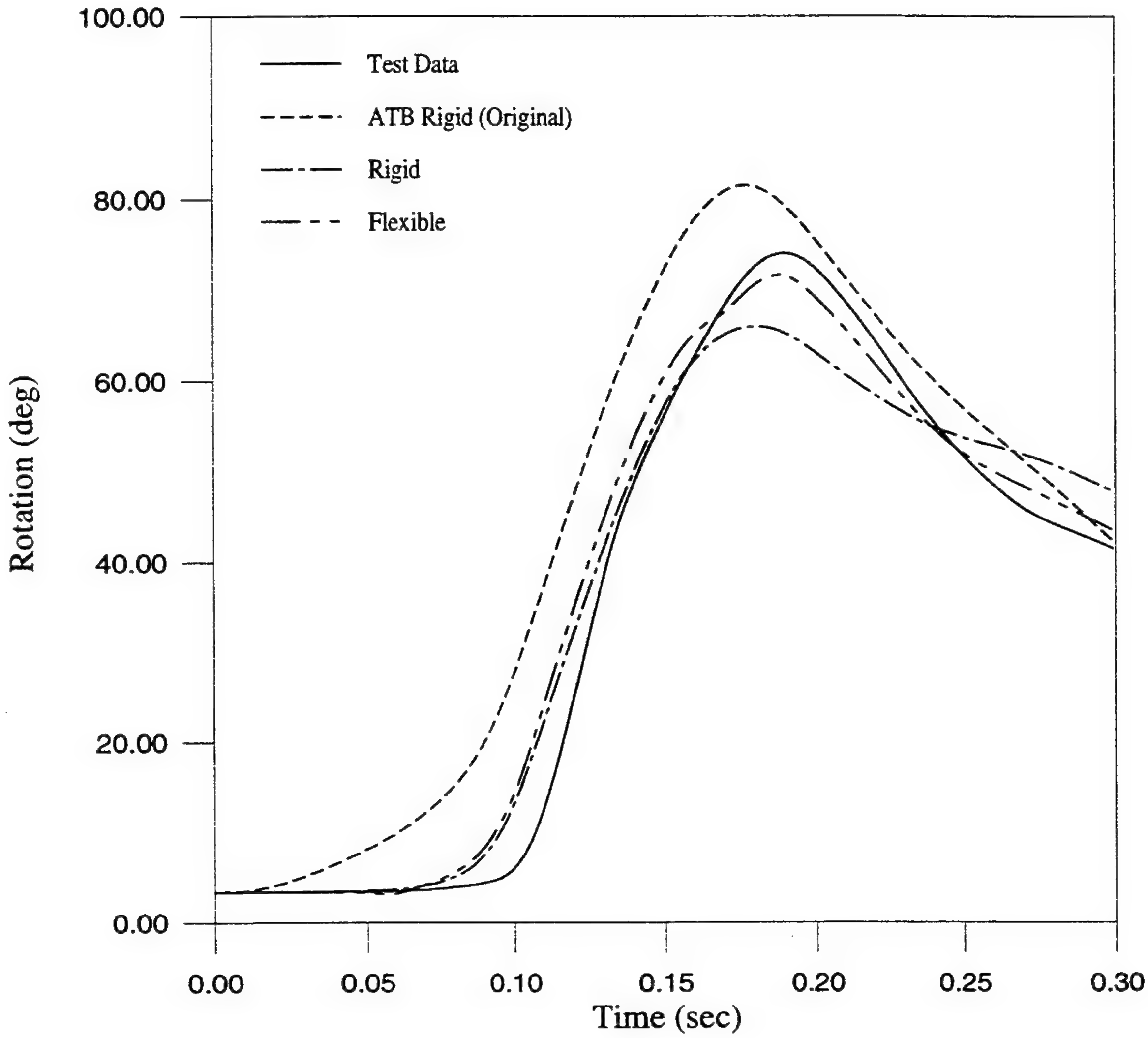


Figure 22. Subject 133 head rotation in the 12g sled test

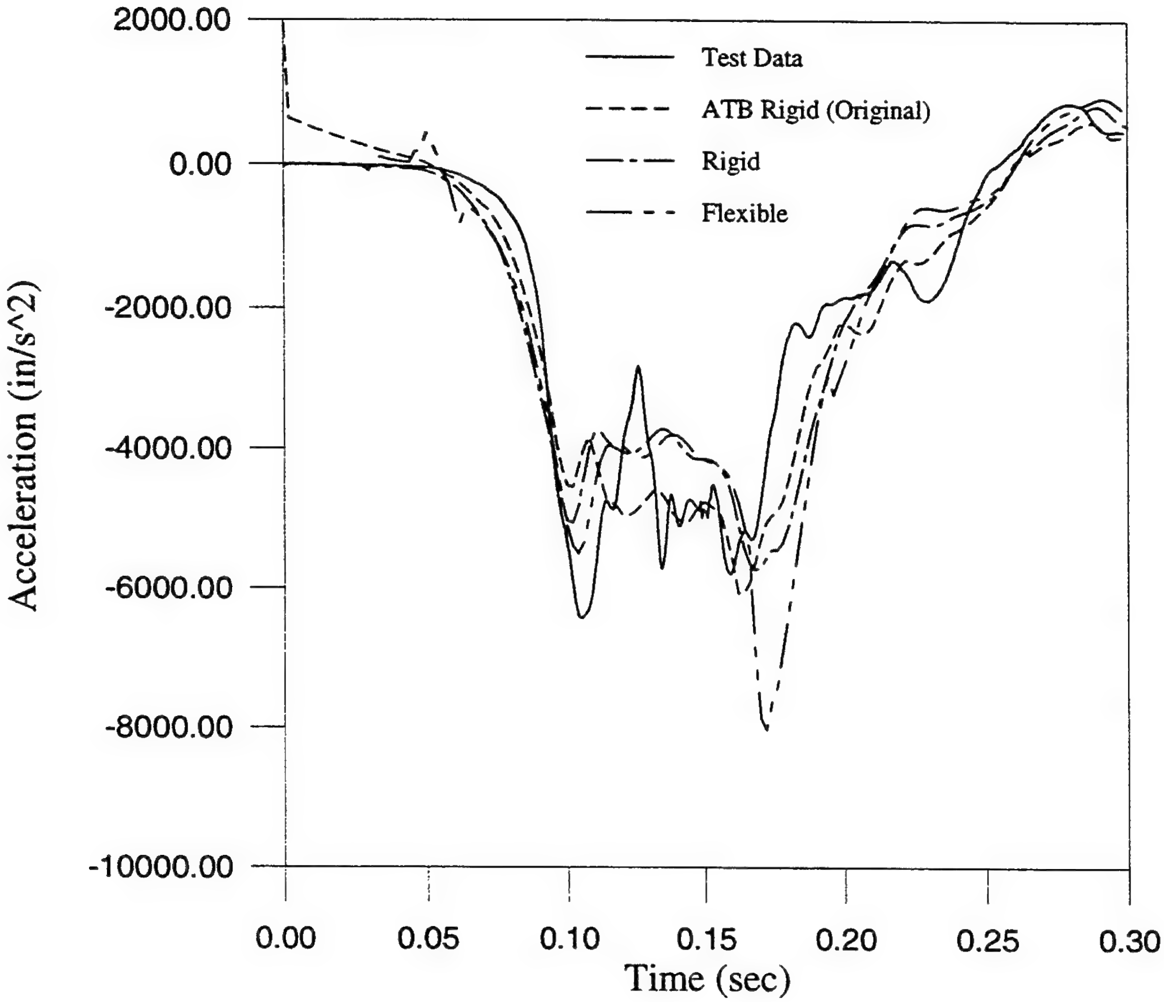


Figure 23. Subject 133 head x-acceleration in the 10g sled test

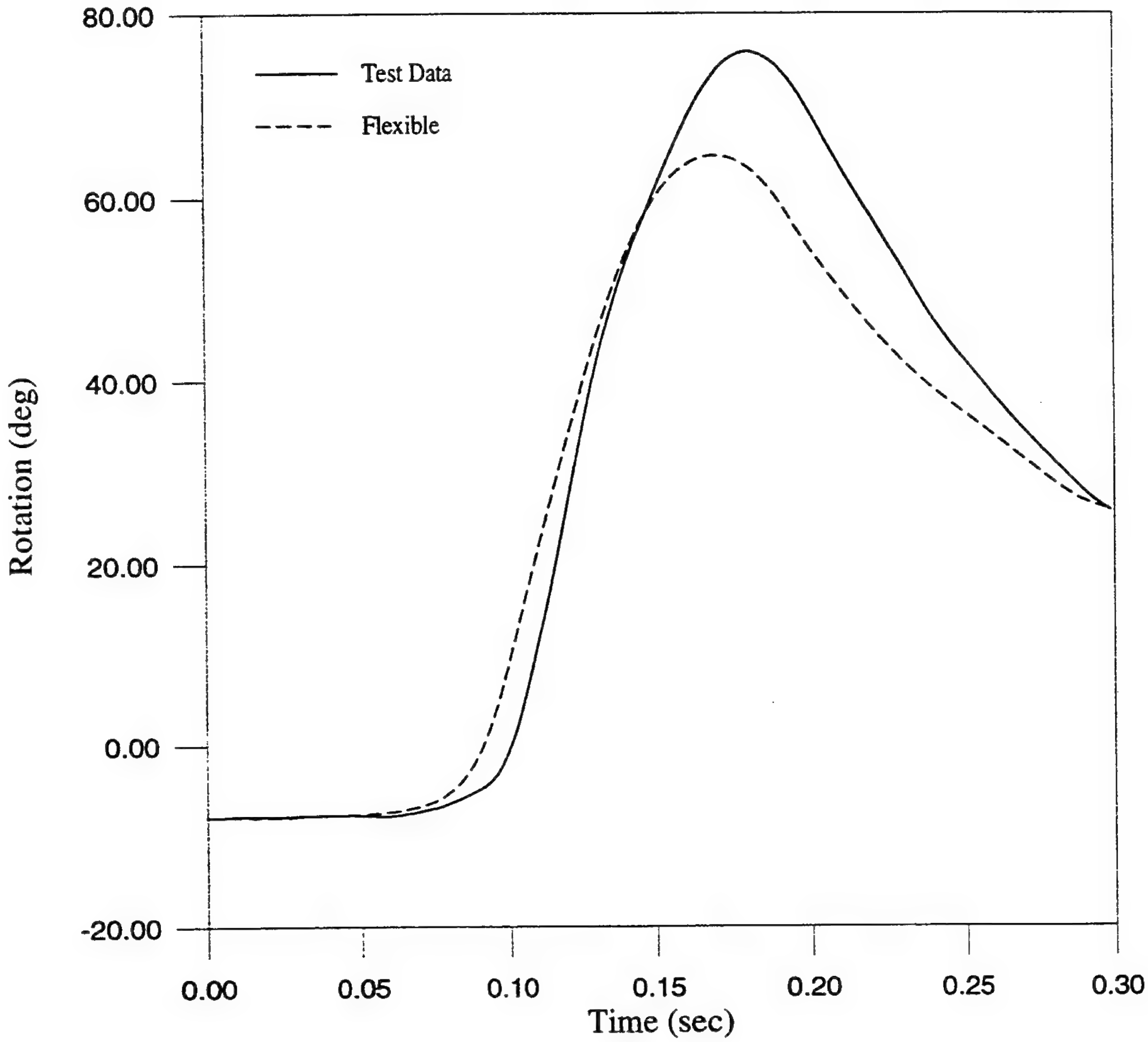


Figure 24. Subject 133 head rotation in the 12g sled test

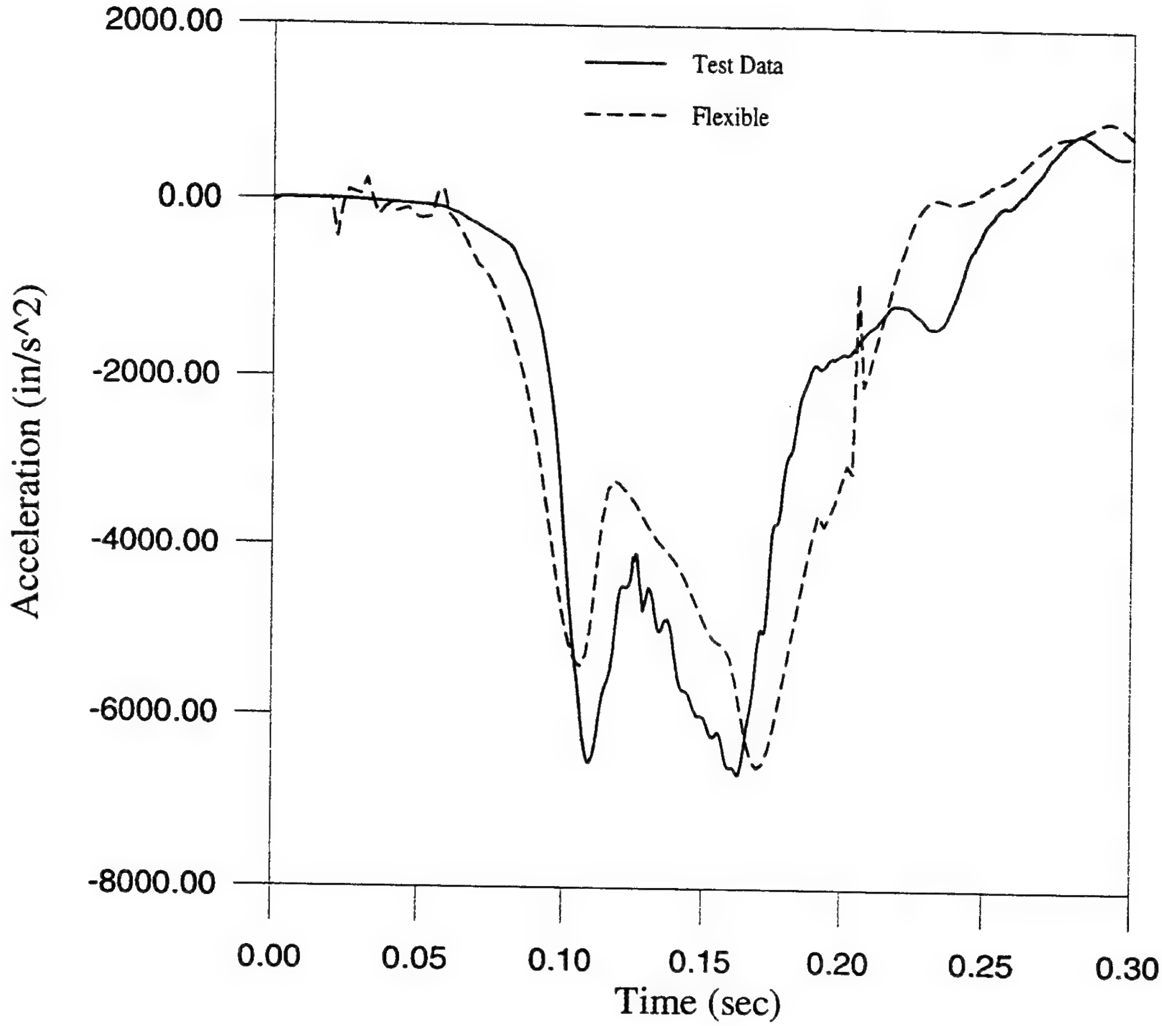


Figure 25. Subject 133 head x-acceleration in the 12g sled test

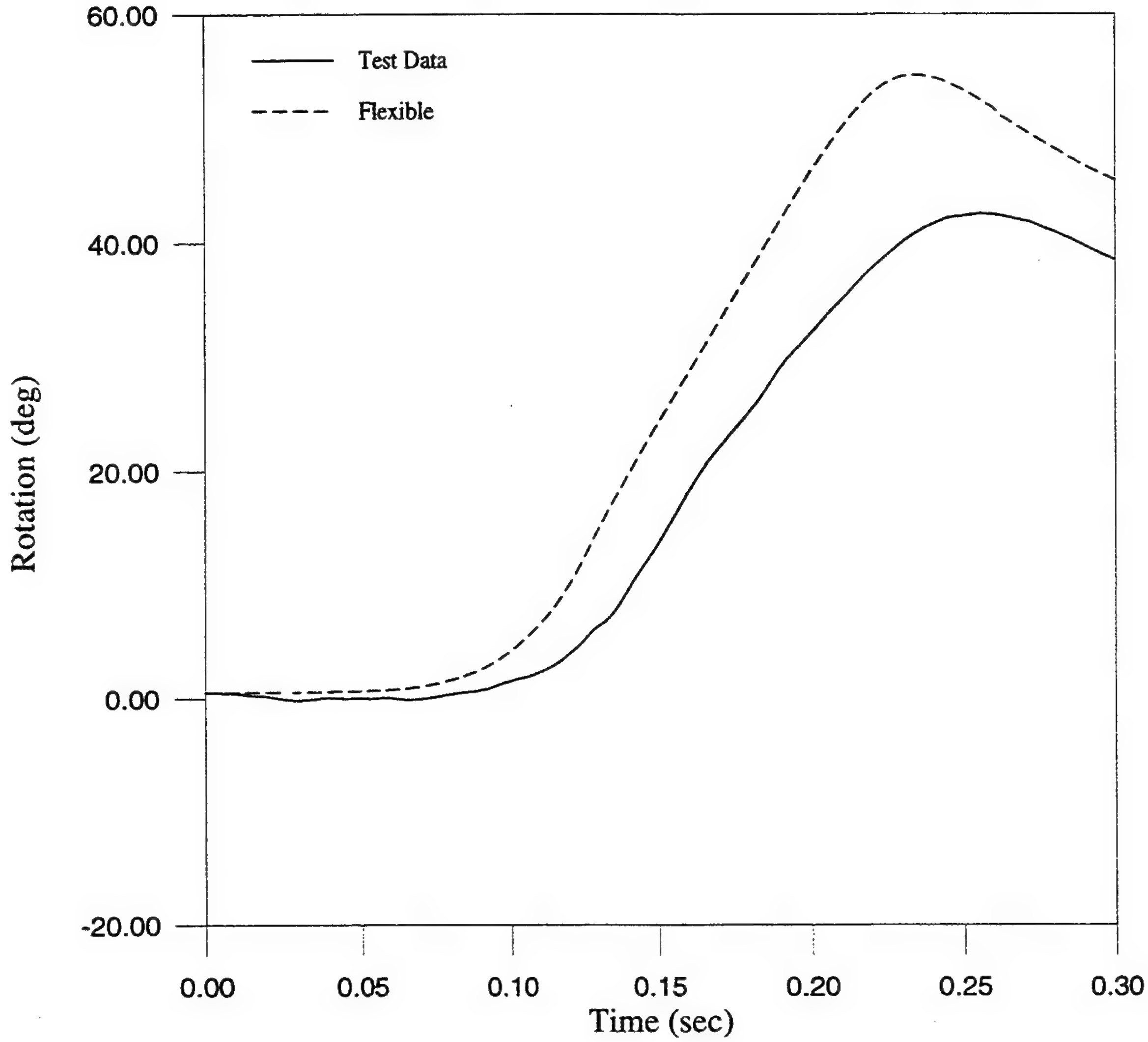


Figure 26. Subject 134 head rotation in the 4g sled test

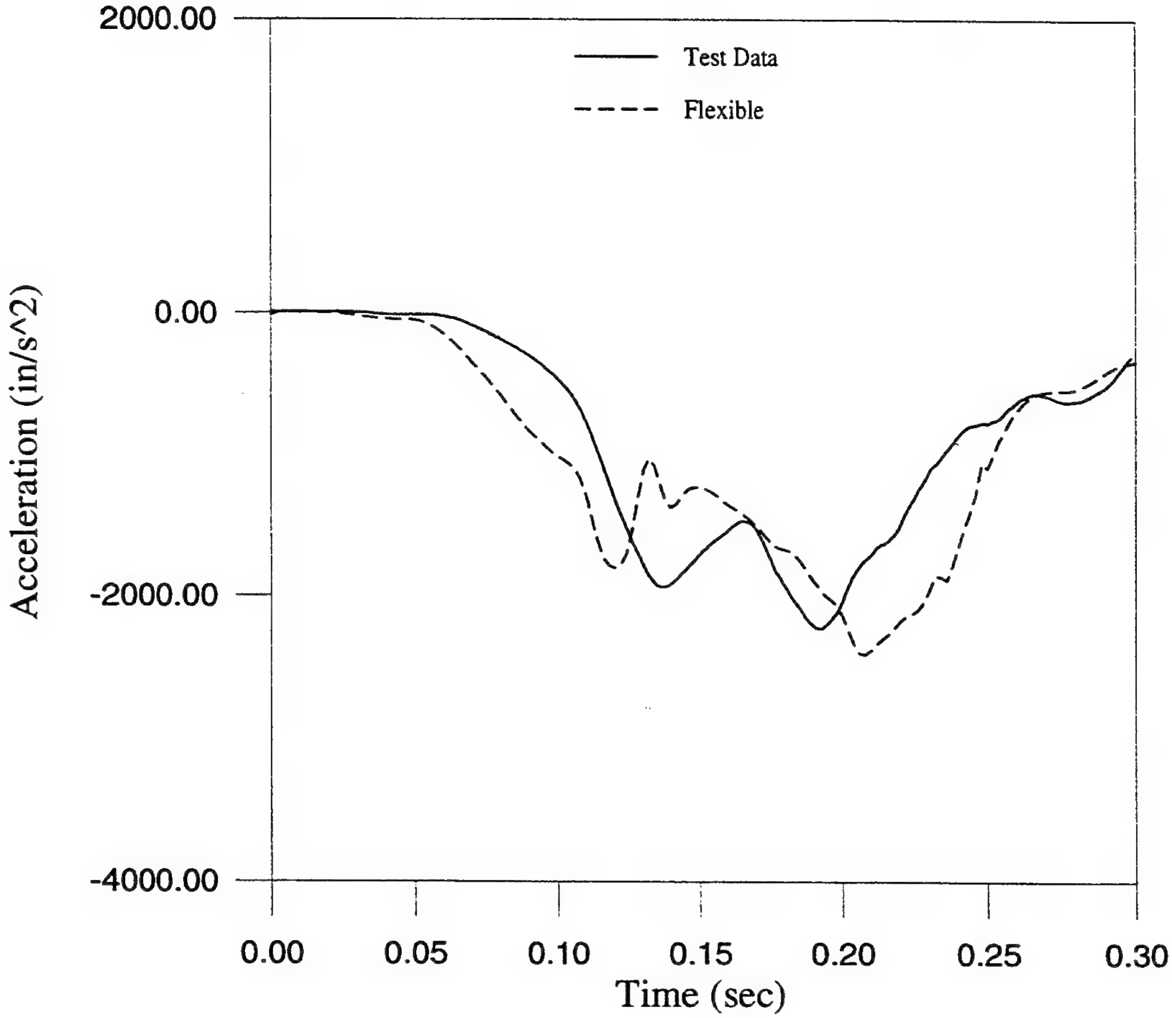


Figure 27. Subject 134 head x-acceleration in the 4g sled test

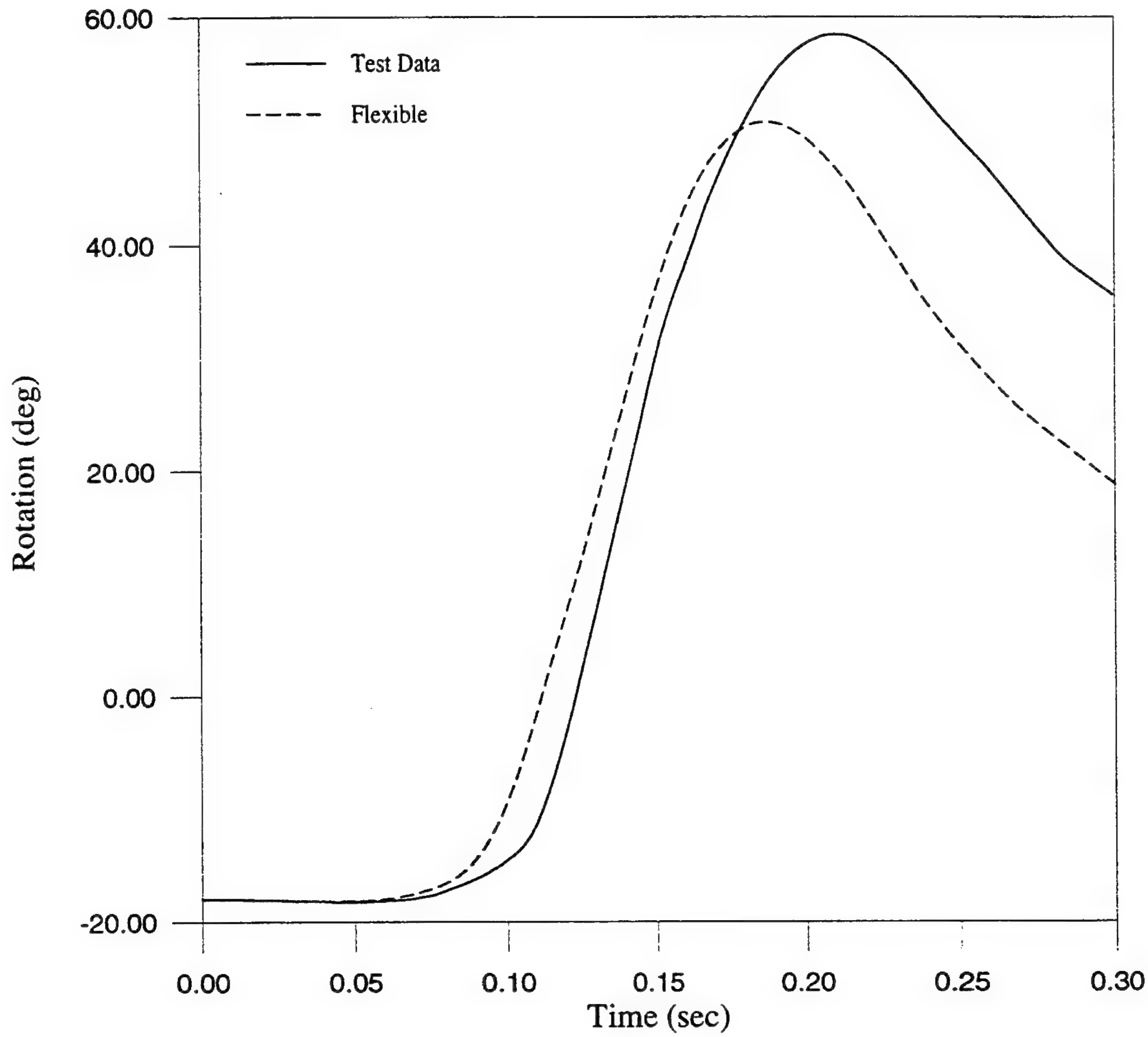


Figure 28. Subject 134 head rotation in the 8g sled test

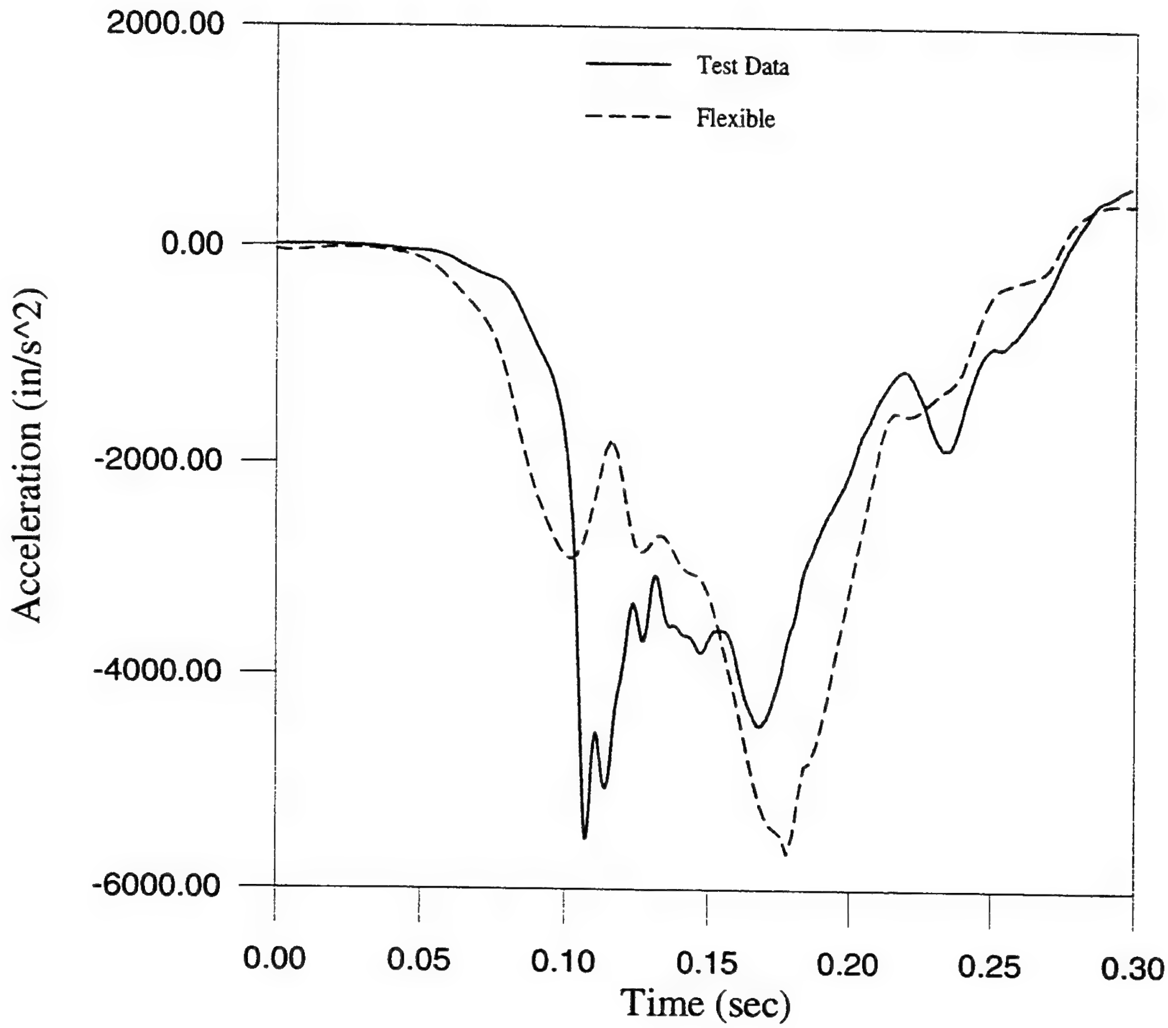


Figure 29. Subject 134 head x-acceleration in the 8g sled test

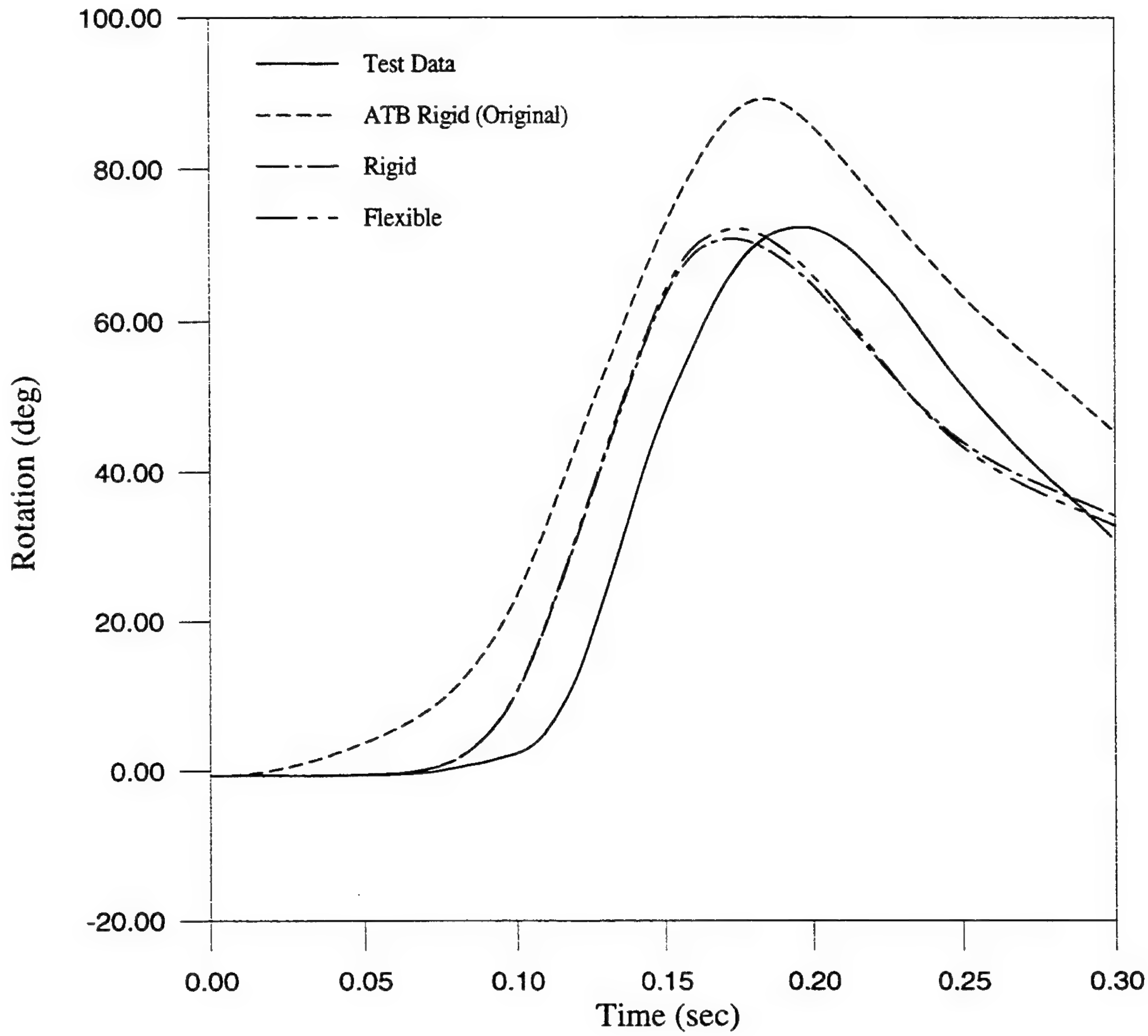


Figure 30. Subject 134 head rotation in the 10g sled test

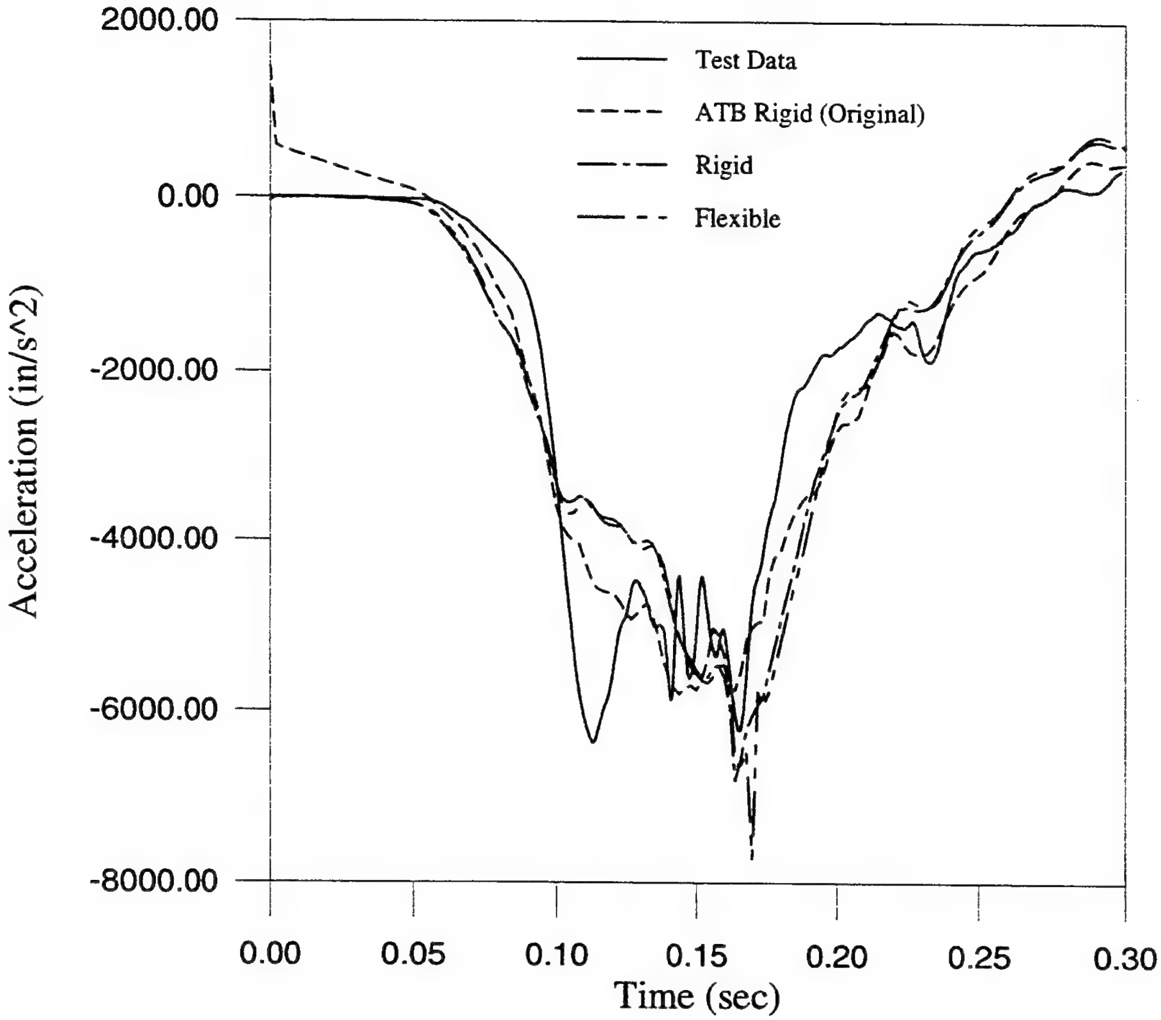


Figure 31. Subject 134 head x-acceleration in the 10g sled test

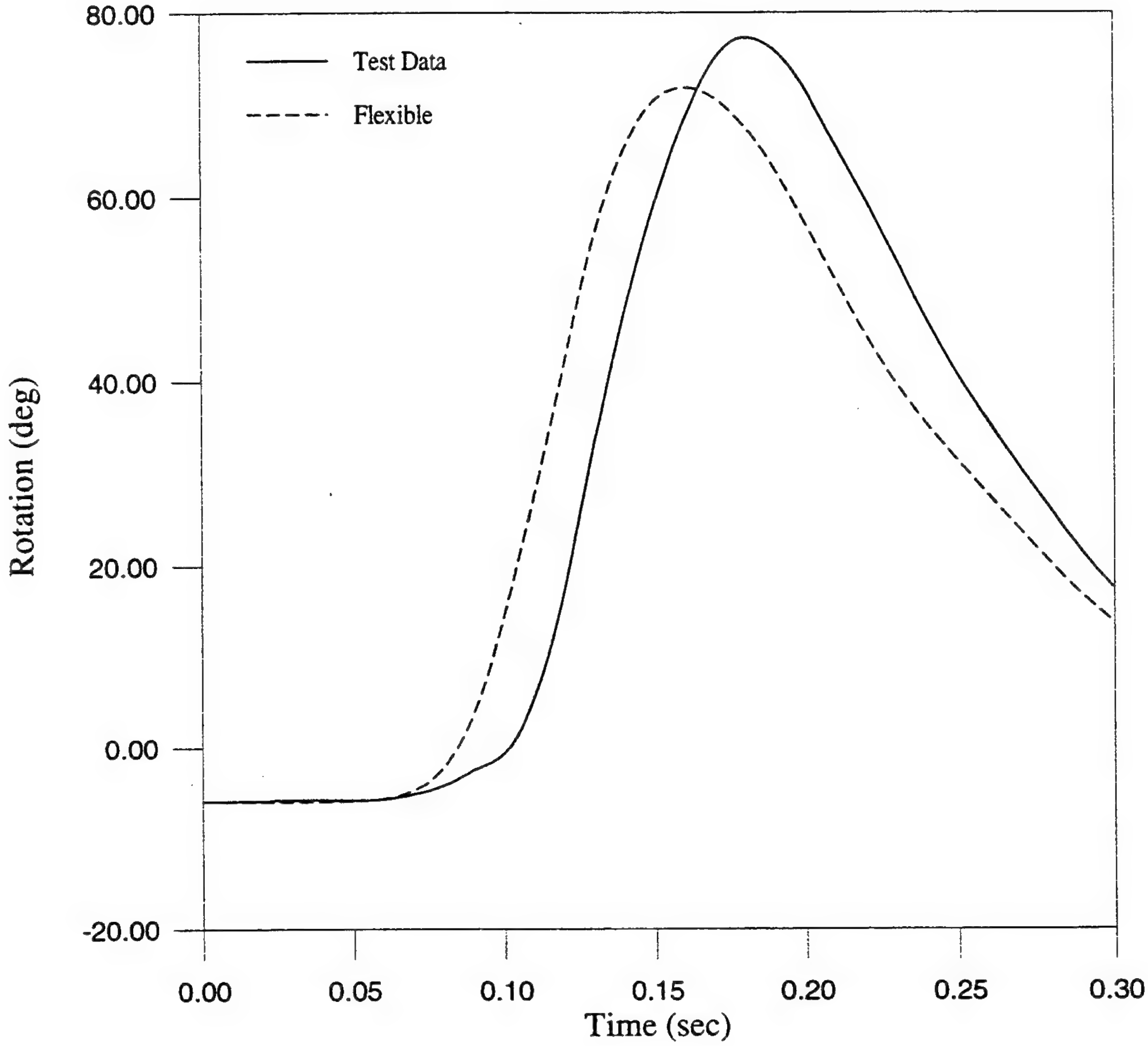


Figure 32. Head Rotation Comparison Subject 134 at 12g.

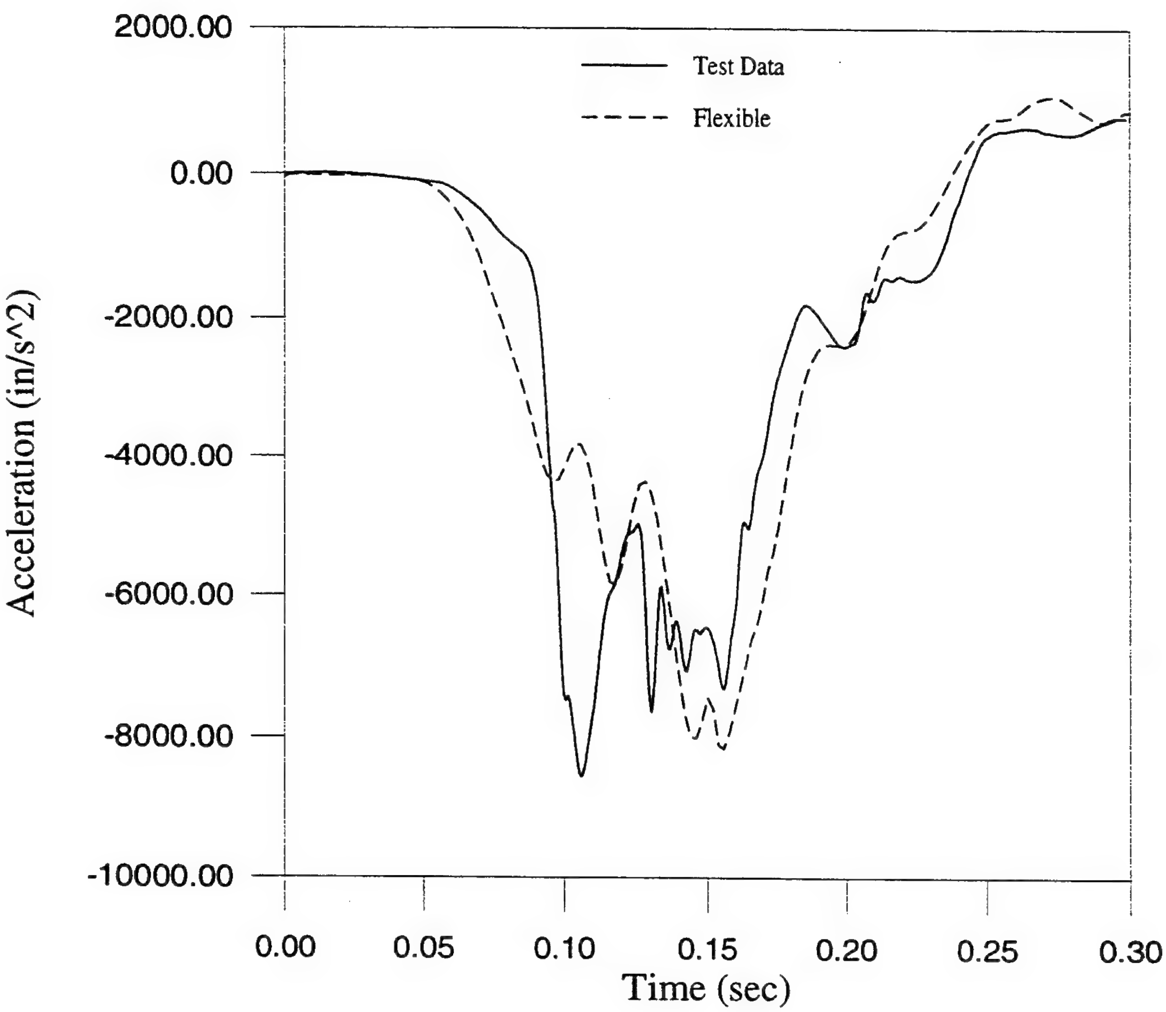


Figure 33. Subject 134 head x-acceleration in the 12g sled test

AN EXAMINATION OF THE VALIDITY OF THE EXPERIMENTAL AIR FORCE ASVAB COMPOSITES

Michael J. Burke
Professor
Department of Psychology

Tulane University
New Orleans, LA 701118

Final Report for:
Summer Research Extension Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

and

Tulane University

December 1995

AN EXAMINATION OF THE VALIDITY OF THE EXPERIMENTAL AIR FORCE ASVAB COMPOSITES

Michael J. Burke
Professor
Department of Psychology
Tulane University

Abstract

A primary objective of the Air Force's Advanced Personnel Testing Program (APT) is to examine the criterion-related validity of new types of computer-administered information processing tests with respect to the prediction of technical training school performance. Related to this objective is the recent development of seven experimental Armed Services Vocational Aptitude Battery (ASVAB) composites designed to differentially predict training school performance in seven job clusters. This study examined the simple and differential validity of the seven experimental ASVAB composites for predicting final training school grades with data from ASVAB forms 11, 12, and 13. These analyses were conducted with (a) uncorrected (observed) criterion-related validity coefficients, (b) validity coefficients corrected for range restriction, and (c) validity coefficients corrected for range restriction and experimental composite unreliability. In addition, this study examined the simple and differential validity of the experimental ASVAB composites subsequent to the assignment of 53 previously unclassified jobs to the experimental composites. Although the experimental ASVAB composites were valid predictors of training school performance in the job clusters for which they were developed, the results generally provided little evidence for the differential validity of the experimental ASVAB composites. Therefore, this investigation offers little support for the usefulness of the seven job clusters for examining the possible differential criterion-related validity of the Air Force's APT information processing tests.

AN EXAMINATION OF THE VALIDITY OF
THE EXPERIMENTAL AIR FORCE ASVAB COMPOSITES

Michael J. Burke

Introduction

Over the last several years, the Air Force has been engaged in research on the development of new types of computer-administered information processing tests (e.g., see Kyllonen, in press; Kyllonen & Christal, 1990). More recently, the Air Force has established a program, the Advanced Personnel Testing (APT) program, to evaluate the utility of computer-administered information processing tests for selection and classification purposes. Because of the large number of Air Force jobs and individuals selected each year into the Air Force, the validation of new types of cognitive tests will require relatively large-scale, multi-year projects such as APT. However, even relatively large-scale individual test validation studies will likely need to consider the extent to which individual study results generalize to the population of Air Force jobs. A particularly important consideration when assessing the generalizability of criterion-related test validation findings for new types of cognitive tests will be how representative the jobs in any particular study are of the population of Air Force jobs. As Burke (1994) noted, concerns with the representativeness of the jobs are particularly important as the number of jobs sampled from the domain of jobs becomes relatively small.

Given that occupational areas or clusters of jobs (e.g., such as Mechanical, Administrative, General, and Electrical) can be developed, the problem of sampling jobs from the domain of Air Force jobs (to ensure representativeness) can be more specifically addressed. That is, in order to generalize to the domain of Air Force jobs, it would be necessary to include in any particular study a representative sample of jobs from each of the occupational areas. Clearly, there should be a balance of jobs (and numbers of individuals in these jobs) so that the different occupational areas are represented in proportion to their importance in the set of occupational areas.

Validity of the ASVAB and ASVAB Composites

Over the last three decades, a number of researchers have examined the validity of cognitive tests for predicting performance in military training schools (e.g., Booth-Kewley, 1982; Earles & Ree, 1992; Hunter, 1984,1985; Maier & Fuchs, 1972; Maier & Grafton, 1983; Maier & Truss, 1983; Ree & Earles, 1991; Sims & Hiatt, 1981). This research has indicated that cognitive tests, in particular parallel forms of the ASVAB, provide a general measure of

cognitive ability that is a useful predictor of performance in all military jobs. In recent military studies, the validity of general cognitive ability has been in the range of .60 to .70. Importantly, Hunter (1984) reported that ASVAB composites for forms 8/9/10 have a high degree of validity in the prediction of training performance for all four branches of the services. Hunter's (1984) analyses covered a total of 190 technical training schools and 103,791 trainees.

A key finding in Hunter's (1984) work with ASVAB forms 8/9/10 was that there was little evidence for differential validity. That is, for most occupational areas, the other composites predicted performance as well as the specific composites constructed for that occupational area. It should be noted that in Hunter's (1984) study, as well as in all previous military studies, there has been evidence for differential validity for clerical work. More specifically, speeded tests make a contribution to the prediction of clerical training performance that is in addition to the validity of general cognitive ability. In addition, there has been some evidence for the differential validity of cognitive tests (i.e., Auto Shop tests) for mechanical work; however, more recent military studies have provided weak evidence for the differential prediction of mechanical work.

Recently, the Air Force has formed seven experimental composites each designed to optimally predict performance in one of seven job clusters. That is, the experimental composites have been designed so as to maximize the potential for differential prediction. The experimental ASVAB composites and typical jobs within a job cluster are listed in Table 1. To the extent that a composite predicts performance to a higher degree in its intended job cluster than in other job clusters, then there would be evidence of differential validity for the new composites. The implication from Hunter's work on forms 8/9/10 of the ASVAB (with respect to previous composites) is that one would not expect the newly formed Air Force ASVAB composites to have a high degree of differential validity. This implication, however, is an empirical question to be addressed.

Insert Table 1 about here

Study Objectives

The primary purpose of the present study is to examine the simple and differential validity of the experimental Air Force ASVAB composites for predicting apprentice school training performance (final school grades) with ASVAB forms 11/12/13. The study will employ the meta-analytic procedures of Raju, Burke, Normand, and Langlois (1991) to estimate the validity of each new ASVAB composite for its own job cluster as well as for each of the remaining six job clusters. These analyses will be conducted with (a) uncorrected (observed)

criterion-related validity coefficients, (b) validity coefficients corrected for range restriction, and (c) validity coefficients corrected for range restriction and experimental composite unreliability. Relative to the validity coefficients corrected only for range restriction, the validity coefficients corrected for range restriction and experimental composite unreliability may suggest composites where APT tests might provide the increases in predictive effectiveness over the present ASVAB composites.

A secondary objective of this study is to examine the simple and differential validity of the experimental ASVAB composites subsequent to the assignment of previously unclassified jobs to the new composites. Given the import of operational validity concerns, previously unclassified jobs will be assigned to the experimental composites based on the highest corrected validity coefficient (i.e., corrected only for range restriction), an examination of the standard errors for each of these corrected validity coefficients, and content-related validity considerations.

Method

Participants

The participants were 88,118 non-prior-service Air Force recruits who were tested with parallel ASVAB Forms 11, 12, and 13 during 1984 to 1988. Only individuals who completed technical training school and who had final school grades were included in the study. The research domain was defined as all apprentice (AF level 3) technical training schools with a minimum of 30 individuals who had completed training. Only schools (jobs) with 30 or more students were considered for the meta-analyses since individual studies would likely not be conducted in smaller schools. This decision resulted in the elimination of 61 schools with 635 students, less than 1% of the population of apprentice school students. The demographic characteristics of the 88,118 participants in the research domain are presented in Table 2.

Insert Table 2 about here

Procedure

Analyses with observed correlations. Initially the mean observed validities of the seven experimental ASVAB composites for predicting final school grades were computed for each of the seven job clusters: Administrative, General I, General II, Electrical, Mechanical, Maintenance, and Electrical/Mechanical II. These validities were based on standard scores for the experimental ASVAB composites. For purposes of this study, only the mean observed validity from each meta-analysis is reported. The complete set of results for the 49 meta-analyses based on observed validity coefficients are reported in Burke and Vaslow (1995).

Analyses with corrections for range restriction. For analyses involving corrections for range restriction, it was necessary to initially compute the standard deviation on each of the experimental ASVAB composites in the research domain of 88,118 students. By pooling the data across the 191 jobs (schools), estimates of the unrestricted standard deviation on each of the composites were made (cf. Glass & Stanley, 1970). That is, the respective composite means, standard deviations, and number of students in each of the 191 schools were employed in estimating the unrestricted composite standard deviations (in the research domain). Then, for each individual study in a meta-analysis, the ratio (i.e., u-ratio) of the standard deviation on a composite in a sample (school) to the unrestricted standard deviation was computed. These u-ratios were employed for making range restriction corrections to each observed validity coefficient in a meta-analysis. For purposes of this study, only the mean unrestricted validity from each meta-analysis is reported. The complete set of results for the 49 meta-analyses based on unrestricted validity coefficients are reported in Burke and Vaslow (1995).

Analyses with corrections for range restriction and composite unreliability. Since item-level data were unavailable, sample-based (i.e., school-based) experimental composite reliabilities were estimated using stratified alpha (Rajaratnam, Cronbach, & Gleser, 1965). In order to use stratified alpha, sample-based ASVAB test reliabilities were estimated with KR-21 (Kuder & Richardson, 1937). KR-21 is calculated from the mean, variance, and number of items on a test. Range restriction corrections were made as discussed above. For the speeded test Coding Speed, reliabilities were based on results presented in Palmer, Hartke, Ree, Welsh, and Valentine (1988). For purposes of this study, only the mean corrected (for range restriction and composite unreliability) validity for each meta-analysis is reported. All mean composite reliabilities were equal to or greater than .91. The complete set of results for the 49 meta-analyses based on range restriction and experimental composite unreliability corrections are reported in Burke and Vaslow (1995). The decision to fix criterion reliability at 1.0 is consistent with previous research examining the criterion-related validity of the ASVAB (cf. Earles & Ree, 1992).

Analyses based on the assignment of previously unclassified jobs to the experimental composites. For these analyses, the unrestricted validities for each experimental composite for each of the 53 previously unclassified jobs (schools) were initially estimated. Based on these unrestricted validity coefficients, their respective standard errors, and content validity considerations, each of the 53 jobs was assigned to an experimental composite. Then, the mean unrestricted validities for each experimental composite and for each job cluster were computed based on the total of 191 jobs (schools). In addition, the mean corrected composite validities (corrected for range restriction and composite unreliability) for each job cluster were recomputed

with the newly assigned jobs included (i.e., based on a total of 191 jobs). The unrestricted validity coefficients and standard errors for each of the seven experimental composites and for each of the 53 previously unassigned jobs are reported in Burke and Vaslow (1995). In addition, the complete set of results for the 98 meta-analyses involving the 191 jobs are reported in Burke and Vaslow (1995).

Results and Discussion

Tables 3 and 4 present the mean observed validities and estimated mean unrestricted validities for the seven experimental ASVAB composites, respectively. Although the validity generalization literature has emphasized the interpretation of operational true validity coefficients (i.e., validity coefficients corrected for range restriction and criterion reliability), the present study will stress results for both mean observed validities in Table 3 and estimated unrestricted validities in Table 4. In addition to examining unrestricted validities, placing emphasis on mean observed validities is consistent with Raju, Burke, and Maurer's (1995) work concerning the interpretation of validity coefficients in applied decision contexts. As indicated in Table 3, each experimental ASVAB composite is a good predictor of technical school performance for its intended job cluster. With the exception of composites D (Electrical) and C (General II), each experimental composite predicts performance for its intended job cluster from slightly better to better than the majority of the other composites for that cluster. For instance, composite E predicts performance for the Mechanical job cluster slightly better to better than five of the other six experimental composites. These results are consistent with previous analyses examining the predictive effectiveness of ASVAB composites for military job clusters (cf. Maier & Truss, 1983).

Given the results in Tables 3 and 4, the prediction of performance in job clusters D (Electrical) and C (General II) with the experimental ASVAB composites should be reconsidered. The mean validity of composite D for the Electrical job cluster is substantially less than the mean validities of composites developed for other job clusters. Likewise, the mean validity of composite C for the General II job cluster is substantially less than the mean validities of composites developed for other job clusters.

Noteworthy, with the exception of composite F (Maintenance), the experimental composites do not predict noticeably better for their intended job cluster than they do for other job clusters. For instance, although composite B (General I) is a good predictor of performance in its job cluster, this composite predicts performance better in 5 of the 6 remaining job clusters (and substantially better in terms of predicting performance in job cluster D (Electrical)). This general pattern of results is maintained for the estimated mean unrestricted validities reported in

Table 4. Together, these results are consistent with previous validity generalization analyses of the ASVAB composites for predicting performance in various branches of the military and military job clusters (cf. Hunter, 1984).

Insert Tables 3 and 4 about here

As discussed above, a secondary objective of this study was to examine the simple and differential validity of the experimental ASVAB composites subsequent to the assignment of previously unclassified jobs to the experimental composites. Fifty three previously unclassified jobs were assigned to the experimental composites/job clusters based on the highest corrected validity coefficient for the job (i.e., corrected only for range restriction) across the seven composites, an examination of the standard errors for each of the corrected validity coefficients, and content-related validity considerations. The estimated mean unrestricted validities for the experimental composites based on data from 191 jobs (schools) are presented in Table 5. As expected, the differences in magnitude between the estimated mean unrestricted validities shown in Table 4 for the 138 jobs and those presented in Table 5 for 191 jobs were very small. These small differences are understandable given that the results based on the 138 jobs have large sample sizes for each job cluster and the increase in total sample size for each job cluster (for the analyses based on 191 jobs) was relatively small.

Insert Table 5 about here

The estimated mean validities corrected for range restriction and experimental composite unreliability are presented in Tables 6 and 7 for the previously assigned jobs and the presently assigned jobs, respectively. In comparison to the mean unrestricted validities in Tables 4 and 5, the respective mean unrestricted and disattenuated validities in Tables 6 and 7 are very similar in magnitude. That is, the mean validities in Tables 6 and 7 are slightly greater than the respective validities in Tables 4 and 5. The small increases in mean validity between these sets of results are due to the relatively high reliabilities for the experimental ASVAB composites. As noted above, the estimated mean reliabilities for the experimental ASVAB composites were equal to or greater than .91. In terms of measurement precision, these latter results suggest that there is little room for the APT tests to improve upon the present experimental ASVAB composites with respect to similar constructs. However, the question of incremental validity of the APT tests in comparison to the ASVAB composites is an empirical question to be addressed.

Insert Tables 6 and 7 about here

Conclusion

In general, the present results are consistent with previous military research which indicates that the ASVAB is a very good predictor of military training school performance. Although the experimental ASVAB composites are valid predictors of performance in the job clusters for which they were developed, there was little evidence for the differential validity of the experimental ASVAB composites. Improvements in the prediction of performance for the General II and Electrical job clusters could be explored through the addition of other ASVAB subtests to the respective C and D experimental composites and/or through examining the possible incremental validity of selected APT tests for these as well as other job clusters/composites. In summary, this investigation provides little support for the usefulness of the seven job clusters for examining the potential differential validity of information processing tests developed within the Air Force's APT program.

References

Booth-Kewley, S. (1983). Predictive validation of the Armed Services Vocational Aptitude Battery forms 8,9, and 10 against performance in 100 Navy schools. San Diego: Navy Personnel Research and Development.

Burke, M.J. (1994). An empirical examination of the effect of second-order sampling error on ASVAB-training proficiency validity estimates. Final report #6 for 1994 Air Force Summer Faculty Research Program. Brooks Air Force Base, TX: Armstrong Laboratory.

Burke, M.J., & Vaslow, J. (1995). Validity Generalization of the Air Force's Experimental ASVAB Composites: Technical Results. Supplemental report to Air Force Office of Scientific Research SREP Subcontract #95-0811. Brooks Air Force Base, TX: Armstrong Laboratory.

Earles, J.A., & Ree, M.J. (1992). The predictive validity of the ASVAB for training grades. Educational and Psychological Measurement, , 721-725.

Glass, G.V., & Stanley, J.C. (1970). Statistical methods in education and psychology. Englewood Cliffs, NJ: Prentice-Hall.

Hunter, J.E. (1984). The validity of the Armed Services Vocational Aptitude Battery (ASVAB) high school composites. Rockville, MD: Research Applications, Inc.

Hunter, J.E. (1985). Differential validity across jobs in the military. Rockville, MD: Research Applications, Inc.

Kuder, G.F., & Richardson, M.W. (1937). The theory of the estimation of test reliability. Psychometrika, 2, 151-160.

Kyllonen, P.C. (in press). Aptitude testing inspired by information processing: A test of the four-sources model. Journal of General Psychology.

Kyllonen, P.C., & Christal, R.E. (1990). Reasoning ability is (little more than) working-memory capacity?! Intelligence, 14, 389-433.

Maier, M.H., & Fuchs, E.F. (1972). Development and evaluation of a new ACB and aptitude area system. Arlington, VA: U.S. Army, Behavior and Systems Laboratory, Technical Research Note 329.

Maier, M.H., & Grafton, F.C. (1981). Aptitude composites for ASVAB 8, 9, and 10. Alexandria, VA: U.S. Army research Institute for the Behavioral and Social Sciences, Research Report 1308.

Maier, M.H., & Truss, A.R. (1983). Validity of ASVAB forms 8, 9, and 10 for Marine Corps training courses: Subtests and current composites. Alexandria, VA: Center for Naval Analyses.

Palmer, P., Hartke, D.D., Ree, M.J., Welsh, J.R., Jr., & Valentine, L.D., Jr. (1988). Armed Services Vocational Aptitude Battery (ASVAB): Alternate forms reliability (Forms 8, 9, 10, and 11). (AFHRL-TP-87-48, AD A191 658). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Rajaratnam, N., Cronbach, L.J., & Gleser, G.C. (1965). Generalizability of stratified-parallel tests. Psychometrika, 30, 39-56.

Raju, N.S., Burke, M.J., & Maurer, T. (1995). A note on direct range restriction corrections in utility analysis. Personnel Psychology, 48, 143-149.

Raju, N.S., Burke, M.J., Normand, J., & Langlois, G.M. (1991). A new meta-analytic approach. Journal of Applied Psychology, 76, 432-446.

Ree, M.J., & Carretta, T.R. (1994). Factor analysis of the ASVAB: Confirming a Vernon-like structure. Educational and Psychological Measurement, 54, 459-463.

Ree, M.J., & Earles, J.A. (1991). Predicting training success: Not much more than g. Personnel Psychology, 44, 321-332.

Sims, W.H., & Hiatt, C.M. (1981). Validation of the Armed Services Vocational Aptitude Battery (ASVAB) forms 6 and 7 with applications to ASVAB forms 8, 9, and 10. Alexandria, VA: Center for Naval Analyses, Marine Corps Operations Analysis Group, CNS 1160.

Table 1

New ASVAB Composites

Experimental Composites/ Job Clusters	Highly Weighted Subtests	Typical Jobs
A (Administrative)	AR, WK, MK	Personnel Spec Accounting Spec
B (General I)	GS, AR, MK, EI	Imagery Interpreter Medical Laboratory Comm Radio Operator
C (General II)	GS, AR, WK, PC, AS, MK	Security Spec Law Enforcement Diet Therapy
D (Electrical)	AR, PC, MK, EI	Electronic Intell Opers Avionic Nav Sys
E (Mechanical)	AR, AS, MK, EI	Vehicle Maint Aero Ground Equipment
F (Maintenance)	GS, AR, WK, CS, AS, MK, MC, EI	Tactical Aircraft Maint Explos. Ordnance-Disposal
G (Electrical/Mechanical II)	WK, PC, CS, AS, MK	Fire Protection Aircraft Armament Syst Fuel System Mechanic

Note. AR (Arithmetic Reasoning), AS (Auto and Shop Information), CS (Coding Speed), EI (Electronics Information), GS (General Science), MC (Mechanical Comprehension), MK (Mathematics Knowledge), PC (Paragraph Comprehension), and WK (Word Knowledge).

Table 2

Demographic Characteristics of Participants in Research Domain of 191 Schools

Category	N	Percent
Gender		
Male	72,422	82.2
Female	15,696	17.8
Race/Ethnic Group		
American Indian	260	.3
Asian	1,572	1.8
Black	12,927	14.7
Hispanic	2,515	2.9
White	70,844	80.4
Educational Level		
Less than High School	745	.9
High School Graduate	69,919	80.2
Some College Exp.	14,799	16.7
Associates Degree	1,208	1.4
College Graduate	1,546	1.8
Age at Entry		
17-18	25,521	29.0
19-20	33,128	37.6
21-22	16,759	19.0
23+	12,710	14.4

Table 3

Mean Observed Validities of the Experimental ASVAB Composites for Seven Job Clusters

			Composite[a]						
Job Cluster	No. of Jobs	N	A	B	C	D	E	F	G
A (Administrative)	21	13,528	**.41**	.36	.28	.37	.33	.35	.39
B (General I)	33	12,885	.44	**.43**	.36	.44	.41	.42	.45
C (General II)	8	16,405	.44	.43	**.37**	.47	.43	.44	.46
D (Electrical)	27	9,816	.48	.51	.44	**.46**	.48	.47	.51
E (Mechanical)	11	5,080	.40	.46	.48	.41	**.48**	.45	.44
F (Maintenance)	6	9,171	.48	.51	.50	.50	.53	**.51**	.51
G (Electrical/Mechanical II)	6	17,343	.42	.43	.42	.43	.45	.45	**.44**

Note. AR (Arithmetic Reasoning), AS (Auto and Shop Information), CS (Coding Speed), EI (Electronics Information), GS (General Science), MC (Mechanical Comprehension), MK (Mathematics Knowledge), PC (Paragraph Comprehension), and WK (Word Knowledge).

[a]A = AR, WK, MK; B = GS, AR, MK, EI; C = GS, AR, WK, PC, AS, MK; D = AR, PC, MK, EI; E = AR, AS, MK, EI; F = GS, AR, WK, CS, AS, MK, MC, EI; G = WK, PC, CS, AS, MK.

Table 4

Estimated Mean Unrestricted Validities of the Experimental ASVAB Composites for Seven Job Clusters

			Composite[a]						
Job Cluster	No. of Jobs	N	A	B	C	D	E	F	G
A (Administrative)	21	13,528	**.46**	.40	.30	.41	.36	.38	.44
B (General I)	33	12,885	.54	**.51**	.41	.52	.47	.48	.54
C (General II)	8	16,405	.51	.49	**.41**	.52	.47	.49	.52
D (Electrical)	27	9,816	.61	.67	.55	**.59**	.60	.57	.65
E (Mechanical)	11	5,080	.44	.53	.59	.48	**.58**	.53	.50
F (Maintenance)	6	9,171	.51	.57	.61	.55	.62	**.61**	.56
G (Electrical/Mechanical II)	32	17,343	.47	.50	.50	.49	.53	.51	**.50**

Note. AR (Arithmetic Reasoning), AS (Auto and Shop Information), CS (Coding Speed), EI (Electronics Information), GS (General Science), MC (Mechanical Comprehension), MK (Mathematics Knowledge), PC (Paragraph Comprehension), and WK (Word Knowledge).

[a]A = AR, WK, MK; B = GS, AR, MK, EI; C = GS, AR, WK, PC, AS, MK; D = AR, PC, MK, EI; E = AR, AS, MK, EI; F = GS, AR, WK, CS, AS, MK, MC, EI; G = WK, PC, CS, AS, MK.

Table 5

Estimated Mean Unrestricted Validities of the Experimental ASVAB Composites for Seven Revised Job Clusters

			Composite[a]						
Revised Job Cluster	No. of Jobs	N	A	B	C	D	E	F	G
A (Administrative)	33	14,859	**.47**	.40	.30	.41	.36	.38	.44
B (General I)	40	13,627	.55	**.52**	.41	.52	.47	.48	.55
C (General II)	8	16,405	.51	.49	**.41**	.52	.47	.49	.52
D (Electrical)	32	10,084	.61	.67	.55	**.59**	.59	.57	.65
E (Mechanical)	18	5,463	.44	.53	.59	.48	**.58**	.53	.50
F (Maintenance)	7	9,255	.51	.57	.61	.55	.62	**.60**	.56
G (Electrical/Mechanical II)	53	18,425	.47	.51	.50	.50	.53	.51	**.51**

Note. AR (Arithmetic Reasoning), AS (Auto and Shop Information), CS (Coding Speed), EI (Electronics Information), GS (General Science), MC (Mechanical Comprehension), MK (Mathematics Knowledge), PC (Paragraph Comprehension), and WK (Word Knowledge).

[a]A = AR, WK, MK; B = GS, AR, MK, EI; C = GS, AR, WK, PC, AS, MK; D = AR, PC, MK, EI; E = AR, AS, MK, EI; F = GS, AR, WK, CS, AS, MK, MC, EI; G = WK, PC, CS, AS, MK.

Table 6

Estimated Mean Unrestricted and Disattenuated Validities of the Experimental ASVAB Composites for Seven Job Clusters

			Composite[a]						
Job Cluster	No. of Jobs	N	A	B	C	D	E	F	G
A (Administrative)	21	13,528	**.48**	.41	.31	.41	.36	.39	.45
B (General I)	33	12,885	.57	**.52**	.42	.53	.48	.50	.56
C (General II)	8	16,405	.53	.50	**.43**	.53	.48	.50	.54
D (Electrical)	27	9,816	.63	.69	.57	**.60**	.61	.59	.68
E (Mechanical)	11	5,080	.46	.55	.62	.49	**.60**	.55	.52
F (Maintenance)	6	9,171	.53	.58	.64	.56	.64	**.62**	.58
G (Electrical/Mechanical II)	32	17,343	.48	.51	.53	.50	.54	.53	**.52**

Note. AR (Arithmetic Reasoning), AS (Auto and Shop Information), CS (Coding Speed), EI (Electronics Information), GS (General Science), MC (Mechanical Comprehension), MK (Mathematics Knowledge), PC (Paragraph Comprehension), and WK (Word Knowledge).

[a]A = AR, WK, MK; B = GS, AR, MK, EI; C = GS, AR, WK, PC, AS, MK; D = AR, PC, MK, EI; E = AR, AS, MK, EI; F = GS, AR, WK, CS, AS, MK, MC, EI; G = WK, PC, CS, AS, MK.

Table 7

Estimated Mean Unrestricted and Disattenuated Validities of the Experimental ASVAB Composites for Seven Revised Job Clusters

			Composite[a]						
Revised Job Cluster	No. of Jobs	N	A	B	C	D	E	F	G
A (Administrative)	33	14,859	**.49**	.41	.32	.42	.37	.40	.46
B (General I)	40	13,627	.57	**.53**	.43	.53	.48	.50	.56
C (General II)	8	16,405	.53	.50	**.43**	.53	.48	.50	.54
D (Electrical)	32	10,084	.63	.68	.57	**.60**	.60	.58	.67
E (Mechanical)	18	5,463	.46	.54	.62	.49	**.59**	.55	.52
F (Maintenance)	7	9,255	.53	.58	.64	.56	.63	**.62**	.58
G (Electrical/Mechanical II)	53	18,425	.49	.52	.52	.50	.54	.53	**.52**

Note. AR (Arithmetic Reasoning), AS (Auto and Shop Information), CS (Coding Speed), EI (Electronics Information), GS (General Science), MC (Mechanical Comprehension), MK (Mathematics Knowledge), PC (Paragraph Comprehension), and WK (Word Knowledge).

[a]A = AR, WK, MK; B = GS, AR, MK, EI; C = GS, AR, WK, PC, AS, MK; D = AR, PC, MK, EI; E = AR, AS, MK, EI; F = GS, AR, WK, CS, AS, MK, MC, EI; G = WK, PC, CS, AS, MK.

FUEL IDENTIFICATION BY NEURAL NETWORK ANALYSIS OF THE RESPONSE OF VAPOR SENSITIVE SENSOR ARRAYS

Paul A. Edwards
Associate Professor
Department of Chemistry

Edinboro University of Pennsylvania
Edinboro, PA 16444

Final Report for:
Summer Research Extension Program
Amstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

December 1995

FUEL IDENTIFICATION BY NEURAL NETWORK ANALYSIS OF THE RESPONSE OF VAPOR SENSITIVE SENSOR ARRAYS

Paul A. Edwards
Associate Professor
Department of Chemistry
Edinboro University of Pennsylvania

Abstract

Neural network analysis of the response of an array of vapor sensitive detectors has been used to identify six different types of aviation fuel. The data set included ninety-six samples of JP-4, JP-5, JP-7, JP-8, JetA and aviation gasoline (AvGas). A sample of each neat fuel was injected into a stream of *breathing air* through an injection port from a gas chromatograph. The aspirated sample was then swept from the injection port to the chamber without separation. In the chamber, the sample was exposed to an array of eight vapor sensitive detectors. The response of each detector was sampled one thousand times at two second intervals, digitized and stored. The response of each sensor was then averaged and stored as the final response or *pattern* of each sample. It was clear from a visual inspection of each of the radar plots that there was a characteristic pattern in the response of the array to five of the six different fuel types. This was confirmed using neural network analysis to study the entire data set. A two step procedure was developed to separate the patterns of all six fuel types into their respective classes. In the first step, fuels were separated into one of five groups: JP-4, JP-5, JP-7, AvGas, or a combined JP-8/JetA group. In the second step, the fuels in the combined group were separated into either JP-8 or JetA groups.

FUEL IDENTIFICATION BY NEURAL NETWORK ANALYSIS OF THE RESPONSE OF VAPOR SENSITIVE SENSOR ARRAYS

Paul A. Edwards

Introduction

As the United States military downsizes, it is imperative that methods be developed for the rapid and reliable identification and monitoring of aviation fuels in soil and water samples (1,2). This research brought together two areas of active research, vapor sensitive detector arrays and artificial neural networks, to address this need.

Arrays of vapor sensitive detectors can be constructed to respond in a measurably different manner to samples of pure compounds and mixtures (3). Such arrays are of practical interest because they can be microprocessor controlled and small enough to be easily portable (4). They have been used to detect and identify a number of hazardous vapors (4,5). A review of chemical sensors by Janata, Josowicz, and DeVaney (6) includes references on the use of sensors for detection of odors, gases of environmental interest and the development of an *electronic nose.*

An array of n detectors will generate n signals, in effect a data vector of length n, which ideally will be unique for each class of aviation fuel. Exposing the array to a set of m samples leads to an m-by-n data matrix that must be analyzed to infer the desired classifications. A number of pattern recognition techniques have been used to affect this analysis, including artificial neural networks (3-6).

Artificial neural networks (ANNs) are computer simulations of biological nervous systems (7). A schematic representation of an ANN is displayed in Figure 1. In general terms, numerical information enters a network through a layer of input neurons or nodes and exits through a layer of output nodes. Information passes from the input to the output layer through a hidden layer (or layers). As information is passed through the layers, numerical weights, biases and transfer functions are applied that adjust the connections between the nodes. Wythoff has written an excellent tutorial on backpropagation neural networks (8).

A recent review of pattern recognition techniques by Brown, *et al* (9) noted "The most novel research in pattern recognition involved work with artificial neural networks." Artificial neural networks have already been applied to the classification of fuels based on gas chromatography data and laser induced fluorescence spectra (10,11). The application of ANNs to classification problems such as this involves at least three steps: training, architecture optimization and validation. Training is the automated process of adjusting the weights and biases in the network so that the output patterns generated by the network match those in the training data set. This process is interconnected with architecture optimization, which is the process of determining the number of input, hidden and output nodes required in the ANN to yield the best classification results. Training and architecture optimization continue until an architecture is found that properly classifies the samples in the training set, and, at least in this study, has a minimum size. Validation is the process of demonstrating the power of the trained network to correctly classify patterns not used in the training process.

The goal of this research was to demonstrate that the response of an array of vapor sensitive detectors could be used to correctly identify aviation fuels. In certain cases, the class of a fuel could be inferred by visually inspecting a radar plot of the pattern of the response of the array. In other cases, neural network analysis techniques were required to correctly classify a fuel sample.

Methodology

Overview of the Experiment: Figure 2. reflects the general set-up of the experiment. Neat samples (2 μL) of fuel were injected into a stream (15 mL/minute) of *breathing air* through the injection port from gas chromatograph (Varian Aerograph Series 1200). Only the injection port of the instrument was used to get the sample into the vapor phase and into the stream of carrier air. The temperature of the injection port was approximately 100°C and the temperature of what would normally be the column chamber was 150°C. The aspirated sample, without separation, was then swept from the injection port into the sample chamber through Teflon tubing and exposed to an array of eight vapor sensitive detectors. The response of each detector was digitized and stored at two second intervals one thousand times. The responses of each sensor were then averaged and stored as the final response or *pattern* of each sample. After data collection was complete, the sample chamber was flushed with a second stream of breathing air, until the response of all the sensors was zero on the analog display meter on the sensor control unit.

Data Set: A set of ninety-six (96) aviation fuels was obtained from The Armstrong Laboratory, Environics Directorate, Tyndall AFB, Florida. The samples were stored in a refrigerator and used as received. The classes of fuels represented in the data set, and the number of samples of each class, are displayed in Table 1.

Table 1: The Neat Fuel Data Set

Category	No. of Samples
JP-4	20
JP-5	12
JP-7	8
JP-8	19
JetA	29
AvGas	8

Chamber: The sensors were mounted in a chamber fabricated from schedule 40 PVC pipe and PVC endcaps. Sample entered the chamber through a Teflon tube mounted on the inside of the top endcap. The inlet tube was pointed directly at the center of a mixing propeller. The mixing propeller was connected to a fan motor by a shaft that passed through the bottom. The arrangement of the sensors in the chamber is displayed in Figure 3. Four small vent holes in the top of the chamber allowed sample and carrier gas to escape. One more tube was mounted through the top of the chamber so an additional stream of compressed air could be used to sweep sample out of the chamber after each run. The approximate volume of the chamber was 20 liters.

Sensors: The sensors and sensor controllers were purchased from International Sensor Technology, Irvine CA. The eight sensors were for benzene, diesel fuel, ethyl benzene, fuel oil, gasoline, JP-4, JP-5, and toluene. The ethyl benzene sensor was set at the factory to respond to concentrations in the range 0 to 200 ppm; the fuel oil sensor was set to respond to concentrations of 0 to 100 ppm. The other six sensors were all set to respond to concentrations in the range of 0 to 1000 ppm. The sensors were controlled by an AG80 multi-channel control unit. The response of each channel was available from the control unit as an analog output signal in the range 0 to 1 volt DC. The control unit also had an analog meter that could be used to monitor visually the progress of a run and the sweeping out of the chamber afterwards.

Data Collection: The analog signal from each of the eight sensors was sampled one thousand times at two second intervals, digitized and stored. An IBM PS/2 model 30 personal computer sampled the analog output of the AG80 sensor control unit through a Model No. CYRDAS 1601 analog-to-digital converter card and two Model SSH/4 sample-and-hold boxes from CyberResearch, Inc., Branford, CT. Each

channel of output from the AG80 was connected to a channel on a sample-and-hold box, and the boxes daisy-chained so that all eight channels of output could be digitized almost simultaneously. The data from each run was stored as a text file in the form of an 8x1000 matrix containing 1000 responses for each of the 8 sensors. The software used to accomplish the data acquisition was the DAS-1600 Test and Control Panel provided with the A/D card running under DOS Version 3.30.

Data Processing: The raw data from the sensors was analyzed using Quattro Pro 1.0 (12) on the same computer used for acquisition. An X-Y plot was generated displaying the 1000 responses of each of the eight sensors. An example of such a plot is displayed in Figure 4. As can be seen in the Figure, the response of each sensor increases in the early part of the run and eventually reaches a maximum. As illustrated in Figure 5, this plot could also be used to diagnose problems such as a bad septum in the injection port or improper mixing of the sample. The response displayed in Figure 4 is much smoother than that displayed in Figure 5. This can be confirmed by examining the standard deviations associated with the average responses. The average and standard deviation of each sensor were calculated over several portions of the response curve, i.e., responses 400 to 1000, 500 to 1000 up to 900 to 1000. The eight average responses over the range with the smallest standard deviations were taken to be the pattern for that sample, and were stored as an 8x1 vector. This preliminary data processing was repeated for each of the 96 fuel samples in the data set leading to an 8x96 data matrix. All subsequent computations were executed on a Gateway2000 P5-90 personal computer running Windows for Workgroups 3.1 (13). Radar plots were generated using Excel Version 5.0C (14). Neural network analysis was accomplished using FIP, a "Fuel Identification Program," by Faruque et al (15). FIP is a suite of MATLAB^{tm}(16) and Neural Network TOOLBOX^{tm}(7) functions for classification of fuels based on gas chromatography data. The calculations reported here were executed using backpropagation methods (8), unless otherwise indicated.

Results and Discussion

This method requires no sample preparation. That is an attractive feature in comparison to the sample preparation required, for example, for certain GC methods (15).

A sample radar plot for each of the six fuel types in the data set showing the responses of all eight sensors is displayed in Figure 6. A radar plot is a convenient method for displaying multidimensional data in a two-dimensional graph. The 8x1 data vector for each sample represents a point in eight-dimensional space. A polygonal figure was drawn for each sample consisting of lines connecting points on eight axes that radiate from the center of the plot similar to spokes on a wheel. The distance of each point from the center of the plot is proportional to the magnitude of the response of the sensor associated with that axis. A point at the center or origin of the figure would represent a response of 0.0 volts for that sensor. A point at the end of the axis on the peripheral of the figure would represent a response of 1.0 volts for that sensor. The radar plots for JP-5, JP-7 and AvGas displayed in Figure 6 appear unique, while those for JP-4, JP-8 and JetA are similar.

The responses of three sensors were removed from consideration after the average responses, standard deviations and pairwise correlations were calculated for the 8x96 data matrix. The response of the benzene sensor was removed because twenty-eight of the ninety-six responses were zero and the average response was essentially zero. The responses of the diesel, gasoline and JP-4 sensor exhibited pairwise correlations greater than 90%. After the responses of the gasoline and JP-4 sensors were removed, all remaining pairwise correlations in the reduced 5x96 data matrix were less than 75%.

A sample radar plot for each of the six fuel types in the data set showing the responses of the remaining five sensors is displayed in Figure 7. With fewer axes, the radar plots for four of the fuels (JP-4, JP-5, JP-7 and AvGas) now appear unique. The continuing similarity of the radar plots for JP-8 and JetA is understandable after examining the specifications of the fuels (17). JetA and JetA-1 are kerosene-based commercial turbine fuels. The main difference in properties is that JetA-1 has a slightly higher maximum freezing point (-47°C) than JetA (-50°C). JetA is the commercial turbine fuel used in the United States while JetA-1 is used primarily outside the U.S. JP-8 is a kerosene-based military turbine fuel that is very

similar to JetA-1. In fact, there is a NATO specification that enables military forces in Europe to purchase fuel certified as JetA-1, and use it after addition of a special additive package. Given the similarity of JetA, JetA-1 and JP-8, it seems reasonable for the responses of the sensor array to be similar.

The initial attempts to train an artificial neural network (ANN) to classify the 96 fuels were not successful. The network architecture used in these initial attempts always included six output nodes, one for each class of fuel. Calculations were executed with a variety of input data matrices ranging from the raw data of seven of the sensors to the first four principle components of the full data matrix. The number of nodes in the hidden layer was also varied from as many as eight to as few as four. Training did not proceed to fewer than two misclassifications during any of these attempts.

A two step process for identifying the fuels was developed after the similarity between JP-8 and JetA was recognized. In the first step, fuels are assigned into the JP-4, JP-5, JP-7, AvGas and a combined JP-8/JetA group. The network architecture used in first step trainings included five input nodes, one for each sensor, and five output nodes, one for each of the listed classes. In the second step, fuels in the combined group are split into either JP-8 or Jet-A groups. The network architecture used in second step trainings included five input nodes, one for each sensor, and two output nodes, one each for JP-8 and JetA.

Trainings were done first with all ninety-six fuels in the training set to minimize the number of hidden nodes in the network. Minimizing the number of hidden nodes in the network is desirable because it minimizes the number of variables involved in the calculations and, thereby minimizing the chance of *superlearning* (18). The process can be monitored by calculating the ratio of samples in the data set to variables for each architecture trained. The goal of minimizing the architecture is to maximize this ratio, at least getting it greater than 1.00. Niebling (19) claims there is minimum number of hidden nodes required for an ANN. He claims the minimum number of *inner units* (n) is related to the number of *identification areas* (k) where $k =< [n(n+1)/2] + 1$. The value of k in step 1 calculations is 5 leading to a value of 4 for n. Thus the architecture of a minimum ANN for this problem would include five input nodes, four nodes in the hidden layer, and five output nodes. There are 49 variables in the minimum ANN. There are 20 weights connecting each of the 5 input nodes to each of the 4 nodes in the hidden layer. There are 20 weights connecting each of the 4 nodes in the hidden layer to each of the 5 nodes in the output layer. There are also 9 biases, one for each of the nodes in the hidden and output layers. With 96 samples in the training set, the ratio of samples to variables would be 96/49 or 1.96. The smallest network that could be trained successfully for step one (i.e., to zero or one misclassification) had 7 nodes in the hidden layer. This network contained 35 weights connecting the input nodes to the hidden layer, 35 weights connecting the hidden layer to the output nodes, and 12 biases. The ratio of samples in step one to variables was 96/72 or 1.33. The smallest network that could be trained successfully for step two (i.e. to zero misclassifications) had 5 nodes in the hidden layer. This network contained 25 weights connecting the input nodes to the hidden layer, 10 weights connecting the hidden layer to the output nodes, and 7 biases. The ratio of samples in step two to variables was 48/42 or 1.14.

After the size of the neural network was minimized, the network was validated using training subset (TSET) and prediction subset (PSET) pairs. The complete data set was divided into these TSET and PSET pairs by randomly moving approximately 10% of each category of fuels from the data set to the prediction subset and using the remaining 90% of the samples as a training subset. Specifically, two TSET/PSET pairs were generated with compositions as indicated in Table 2. Nine samples (two JP-4, one JP-5, one JP-7, four JP-8 and JetA's, and one AvGas) were moved from the data set to a prediction subset leaving eighty-seven samples in the training subsets. The network was trained with the training subsets using the minimum architecture that could be trained successfully by the complete data set. That architecture had 5 input, 7 hidden and 5 output nodes in the first step, and 5 input, 5 hidden and 2 output nodes in the second step. The trained network was then used to predict the classification of the samples in the prediction subset.

Table 2: Composition of Training and Prediction Subsets

Category	*No. in Training Subset*	*No. in Prediction Subset*
JP-4	18	2
JP-5	11	1
JP-7	7	1
JP-8 and JetA	44	4
AvGas	7	1
Total	87	9

Table 3: Validation Results

TSET/PSET Pair No.	*Training No. Correct Step 1*	*Step 2*	*Prediction No. Correct Step 1*	*Step 2*
1	86/87		8/9	
1	87/87	44/44	7/9	2/4
2	86/87	44/44	9/9	4/4

The results of the validation process are collected in Table 3. The results for TSET/PSET pair number one in step 1 illustrate the dilemma of overtraining. When the network was trained to 100% correct classification, the power to make predictions was less than when training was terminated with one misclassification. The results for TSET/PSET pair number 1 in step 2 are not as good as desirable. The results for TSET/PSET pair number two in both steps are good. The validation process was repeated using radial basis (7) instead of backpropagation methods and those results are collected in Table 4. Only prediction results can be compared because the radial basis method requires training to continue until 100% classification is reached. Interestingly essentially the same results were obtained using the two training methods.

Table 4: Validation Results Using Radial Basis Methods

TSET/PSET Pair No.	*Prediction No. Correct Step 1*	*Step 2*
1	7/9	2/4
2	9/9	4/4

Obviously a number of additional TSET/PSET pairs could be generated. However, we believe the following comments are reasonable. The network is able to correctly classify the fuels and make predictions, *but just barely*. The results of the neural network analysis would probably be better if the responses of more than five sensors were used in the classification process. The fact that only five sensors were used in the classification process may be the result of the decision to use commercially available sensors as opposed to making them ourselves. For example, we were not aware until after data collection had begun that certain sensors sold for different applications were actually physically the same only adjusted or tuned differently. That could, of course, explain the high correlations between the responses of three of the sensors. We were also disappointed by the non-specific response of the sensors. We expected the sensors to respond to compounds that were similar to those for which they were designed. However,

the toluene sensor, for example, responded to a large number of substances. For example, it responded to acetone, the vapors from wax-stripper being used by housekeeping personnel, and perfume!

Conclusions

The results of this study demonstrate that:

- an array of gas sensitive detectors can be constructed so that it responses uniquely to different types of aviation fuels.
- neural network techniques can be used to assist in classifying the response of an array of gas sensitive detectors.
- it is feasible to identify aviation fuels using neural network analysis techniques to analyze the response of an array of vapor sensitive detectors.

Future research should focus on two issues: selecting more appropriate sensors and miniaturizing the array. In this first study, the responses of only five of the eight sensors were ultimately involved in the classification. One sensor yielded too many zero responses to be statistically significant. The responses of three other sensors were highly correlated, so the responses of two were removed from the data matrix. Additional sensors should be found to help in the classification and to expand the number of identifiable classes. The commercial sensors used in the first study came mounted in large, explosion proof housings. These housings were then mounted in a chamber approximately 20 liters in volume. A much smaller housing for the array is needed so that it could be put into a probe and dropped into a test boring. Given that no sample preparation is required, such a field application can be envisioned.

Acknowledgments

This research would not have been completed successfully without the commitment and dedication of three Edinboro University of Pennsylvania **undergraduate** students: Ms. Lynn A. Gilliland, Mr. Craig W. McCarrick, and Mr. David T. Ohmer. PAE will always remember days when: he was told "I told you that days ago", they took off to do something before he thought of it, and they threw him out of his office so they could work! Mr. Donald McCaslin, McCaslin Construction, McKean, PA, constructed the sample chamber. The Institute of Research and Community Service of Edinboro University of Pennsylvania provided matching funds. Dr. Howard Mayfield, Armstrong Laboratory, Environics Directorate, Tyndall AFB, Florida, provided the fuel samples and encouragement throughout the project. Finally, this research was supported by Contract Number F49620-93-C-0063 between Edinboro University of Pennsylvania and the United States Air Force.

Bibliography

1. Lavine, B.K.; Qin, X.; Stine, A.; Mayfield, H.T. *Process Control Quality,* **1992**, *2*, 347-355.

2. Lavine, B.K.; Stine, A.; Mayfield, H.T. *Anal. Chimica Acta.* **1993**, *277*, 357-367.

3. Stetter J.R. Electrochemical Sensors, Sensor Arrays, and Computer Algorithms for Detection and Identification of Airborne Chemicals, In: *Fundamentals and Application of Chemical Sensors:* Schuetzle, D; Hammerle, R.; Ed. ACS Symposium Series, 300, American Chemical Society, Washington, DC. 1986, p. 299-308.

4. Stetter, J.R.; Jurs, P.C.; Rose, S.L. *Anal. Chem.* **1986**, *58*, 860-880.

5. Rose-Pehrsson, S.L.; Grate, J.W.; Ballantine, D.S.;Jr.; Jurs, P.C. *Anal. Chem.* **1988**, *60*, 2801-2811.

6. Janata, J.; Josowicz, M.; Devaney, D.M. *Anal. Chem.* **1994**, *66*, 207R-228R.

7. Neural Network TOOLBOX: For Use with MATLAB^{tm}, The Mathworks, Inc. Natick Massachusetts, 1992.

8. Wythoff, B.J. *Chemom. Intell. Lab Syst.* **1993**, *18*, 115-155.

9. Brown, S.D.; Blank, T.B.; Sum, S.T.; Weyer, L.G. *Anal. Chem.* **1994**, *66*, 315R-359R.

10. Long, J.R.; Mayfield, H.T.; Henley, M.V.; Kromanm, P.R. *Anal. Chem.* **1991**, *63*, 1261-1264.

11. Andrews, J.M.; Lieberman, S.H. *Anal. Chim. Acta.* **1994**, *285*, 237-246.

12. Quattro Pro 1.0, Borland International, Scotts Valley, California, 1989.

13. Windows for Workgroups Version 3.1, Microsoft Corporation, 1994.

14. Excel Version 5.0C. Microsoft Corporation, 1985-1994.

15. Faruque, A.; Mayfield, H.T.; Lavine, B.K.; Edwards, P.A. "Application of Neural Networks to the Classification of Chromatographic Data", poster presentation at Pittcon'95, New Orleans, LA, March. 1995.

16. MATLAB^{tm}, High-Performance Numerical Computation and Visualization Software, The Mathworks, Inc. Natick, Massachusetts, 1992.

17. In: *Handbook of Aviation Fuel Properties.* Coordinating Research Council, Inc., Atlanta, GA, 1983, p. 2-18. (Available from: Society of Automotive Engineers, Inc., General Publications Department, 400 Commonwealth Drive, Warrendale, PA, 15096).

18. Duin, R. P. W. *Pattern Recognit. Lett.* **1994**, *15*, 215-217.

19. Niebling, G. *Sens. Actuators,* **B 1987**, *18-19*, 259-263.

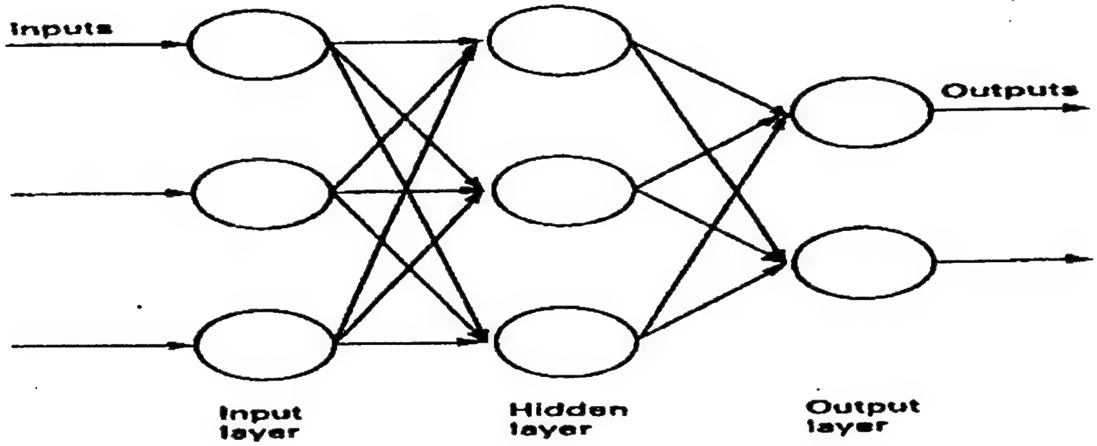


Figure 1: Schematic presentation of an artificial neural network.

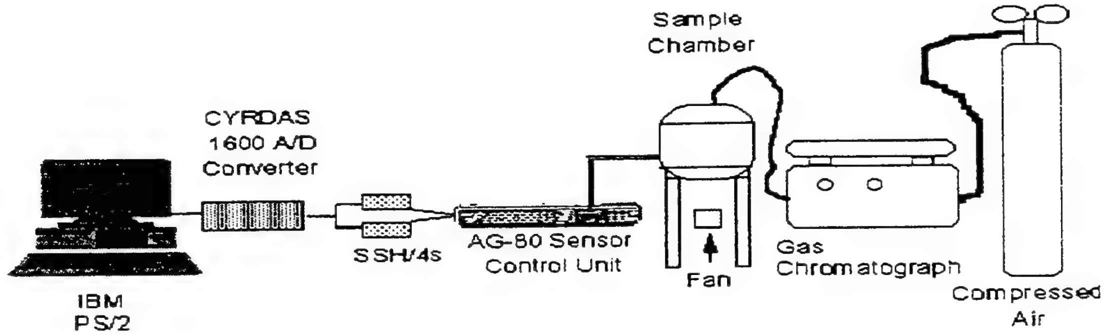


Figure 2: General set-up of the experiment.

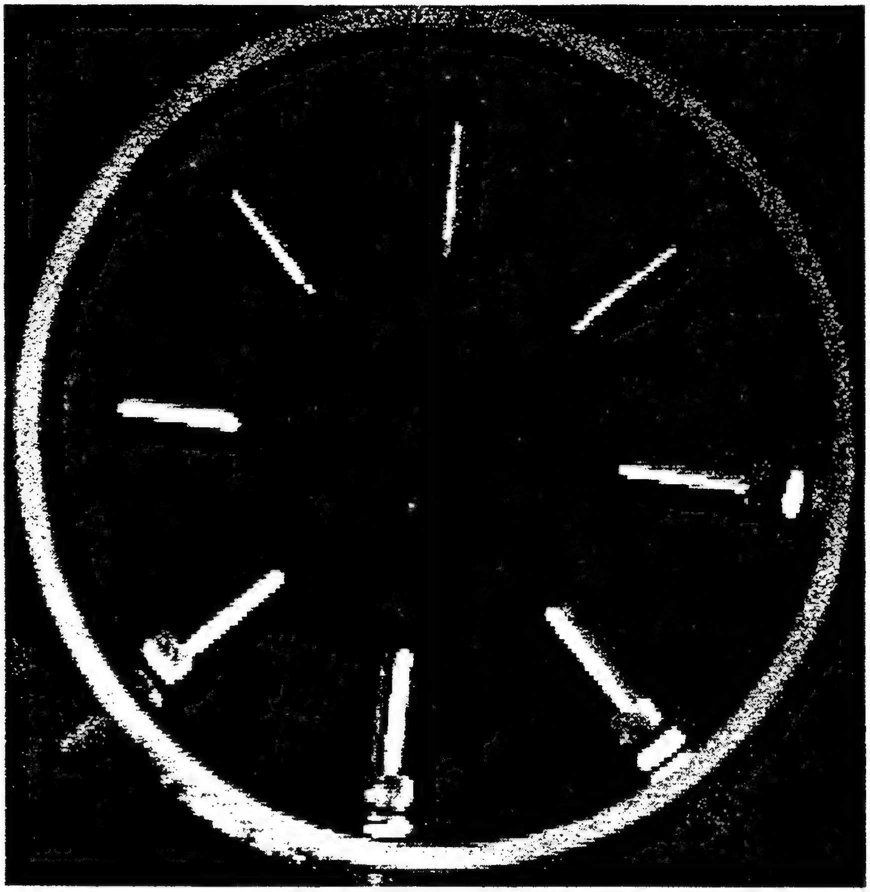


Figure 3: Orientation of the sensors in the sample chamber.

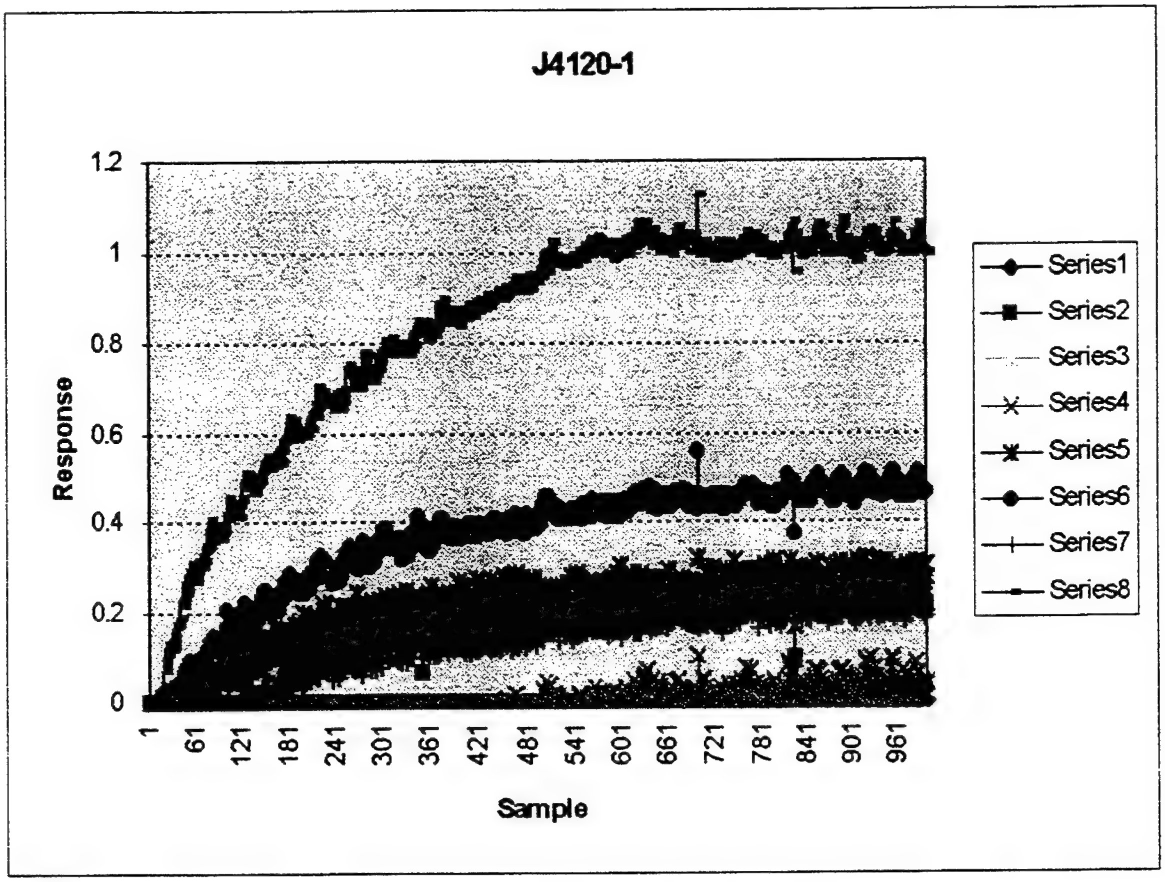


Figure 4. Typical response of the array to a sample as a function of time.

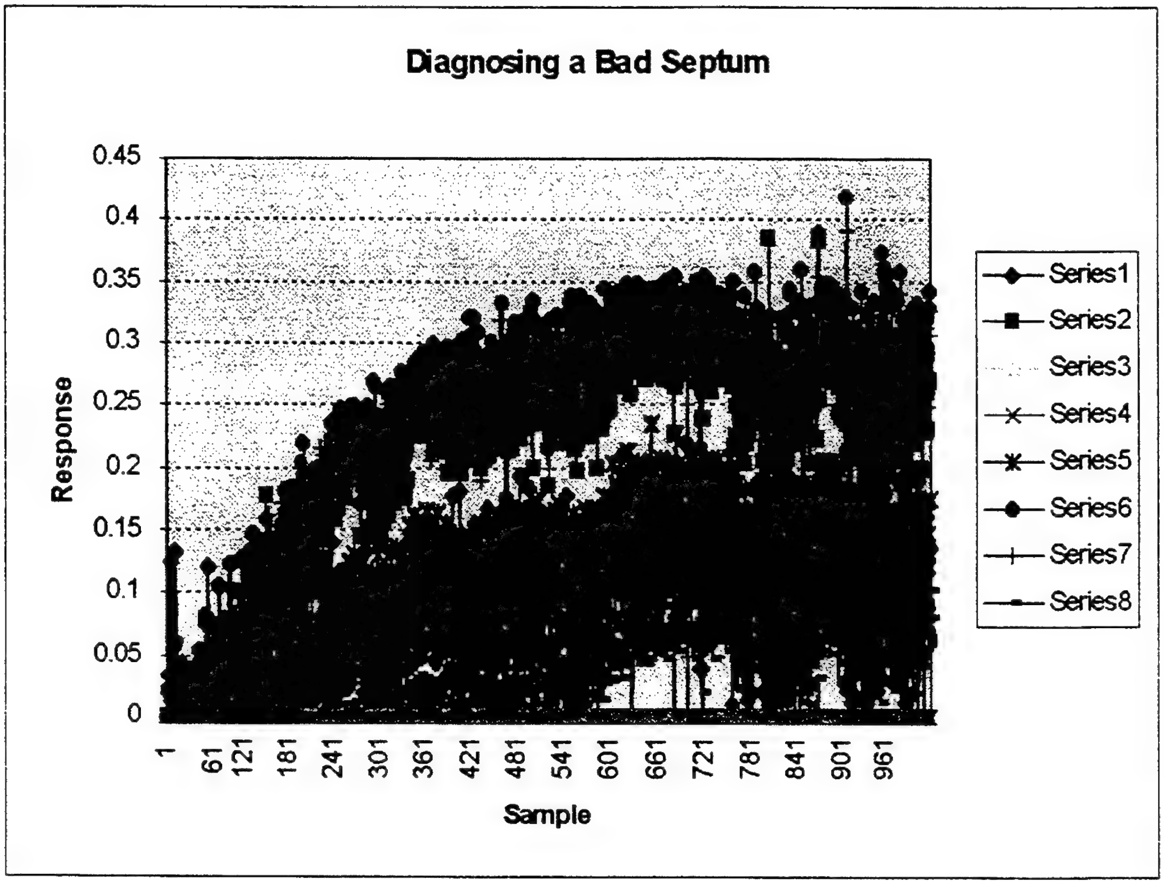


Figure 5. Response of the array when the septum in the injection port is leaking.

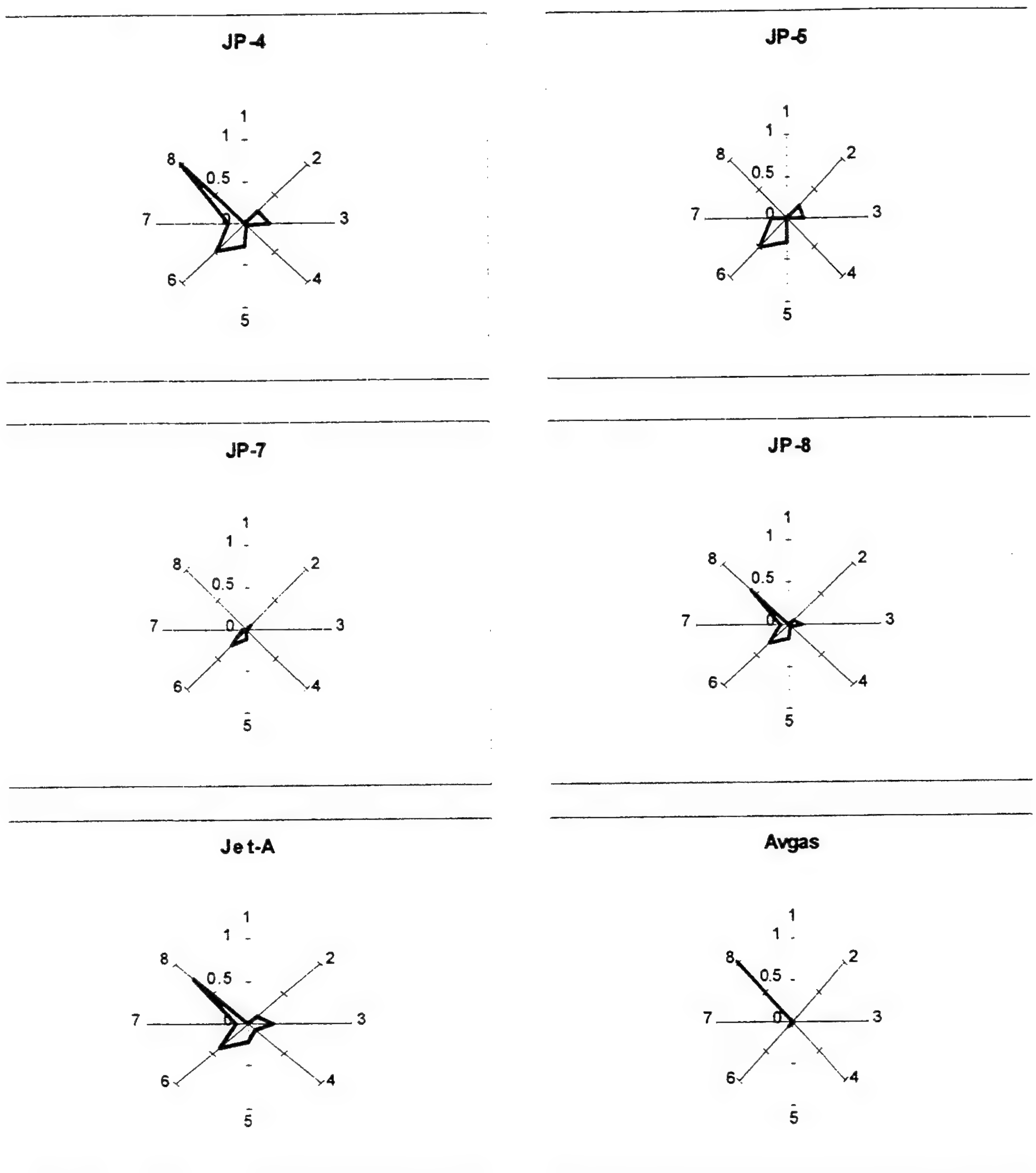


Figure 6: Sample radar plots displaying the response of eight sensors for each fuel category.

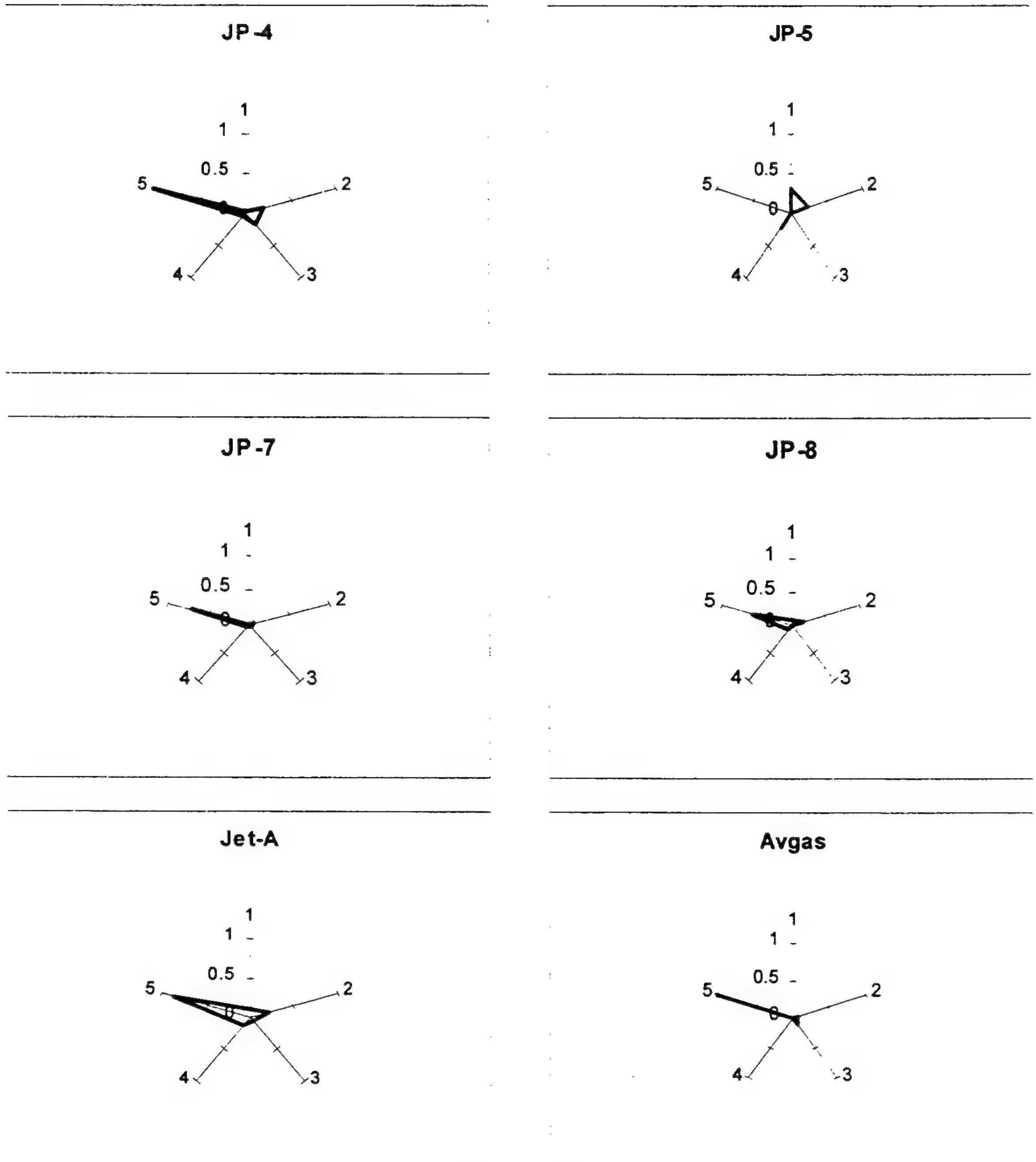


Figure 7: Sample radar plots displaying the responses of five sensors for each fuel category.

A Comparison of Multistep vs. Singlestep Arrhenius Integral Models for Describing Laser Induced Thermal Damage to the Retina

Bernard S. Gerstman
Associate Professor
Department of Physics

Florida International University
Miami, FL

Final Report for:
Summer Research Extension Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

and

Armstrong Laboratory

December 1995

A COMPARISON OF MULTISTEP VS. SINGLESTEP ARRHENIUS INTEGRAL MODELS FOR DESCRIBING LASER INDUCED THERMAL DAMAGE TO THE RETINA

Bernard S. Gerstman
Associate Professor
Department of Physics
Florida International University

Abstract

Laser induced thermal damage to the retina is investigated theoretically. Possible reaction mechanisms and the resulting kinetics are analyzed in terms of their potential relevance to the damage process. The one step Arrhenius type thermal damage integral of Henriques is analyzed for its strengths and weaknesses and additional steps to the reaction mechanism are proposed. However, the zero-order Henriques mechanism is shown to well represent the data for pulse durations greater than 10 microseconds. A zero-order biochemical damage mechanism involving free radical formation in melanosomes and thermal disruption of the melanosomes' membranes is proposed.

A COMPARISON OF MULTISTEP VS. SINGLESTEP ARRHENIUS INTEGRAL MODELS FOR DESCRIBING LASER INDUCED THERMAL DAMAGE TO THE RETINA

Bernard S. Gerstman

I. Introduction

Damage to the visual system caused by lasers has been occurring for decades. Laser interactions with the visual system are also used for beneficial purposes such as laser surgery. As lasers continue to become more powerful, both the danger and the benefits increase. In order to better protect against damage, and to take additional advantage of surgical methods, it is extremely helpful if the underlying biochemical and biophysical effects of laser interactions with the visual system are well understood. In this report we discuss an approach to investigating the biochemical reactions that are caused in the retina as a result of laser energy deposition.

The effectiveness of lasers over conventional sources for surgery and as a source of damage is due to their high fluence (J/cm^2) and small spot size. A laser spot that is millimeters in diameter with a fluence of 10^{-5} J/cm^2 on the cornea is focused by the cornea and lens to micrometer size with fluences of 1 J/cm^2 on the retina. Empirical studies of the fluence as a function of pulse duration laser required to cause damage is summarized in Fig. 1. The data points were obtained by different researchers using different types and wavelengths of lasers and on eyes from different species of animals so care must be used when interpreting the data. However, general trends are clearly discernible.

For pulse lengths longer than a microsecond, a minimal visible lesion requires higher fluence as the pulse duration increases. This has been interpreted as due to a mechanism in which temperature rises cause thermal damage. Since heat is conducted away in the retina on microsecond time scales, a longer pulse will allow more heat to conduct away during deposition which will prevent the irradiated area from reaching a high temperature unless the fluence is increased.

For pulse lengths between approximately a microsecond and a nanosecond, the leveling off of the fluence required to cause damage is due to the duration of the pulse being shorter than heat conduction times away from the irradiated area. In this pulse length regime, all of the deposited energy remains in the irradiated part of the retina during the pulse duration and therefore the energy density in the retinal tissue which causes the damage is independent of the pulse duration. In this pulse length regime, studies have been performed {1,2} that show that bubble production around the RPE is the cause of threshold MVL damage. Above threshold levels, thermal damage may also occur.

For pulse lengths below a nanosecond, the MVL threshold fluence appears to drop. Even more so than with the threshold bubble damage regime, the pulse lengths are significantly shorter than thermal conduction times so all the deposited energy remains localized during deposition. The sub-nanosecond pulses are also shorter than another critical relaxation time which is the time for sound to travel across the absorbing melanosomes which are approximately 1 μm in diameter. This pulse length regime is known as the stress confinement regime because pressure waves traveling at the speed of sound cannot escape during the duration of the pulse. This can lead to the build up of extremely high pressure waves that are transmitted throughout the cell causing damage or to the explosion of the melanosome within the cell. The high pressures generated by stress confinement could be the mechanism for threshold MVL damage and explain why the threshold fluence drops for pulse lengths below a nanosecond. For pulse lengths shorter than a picosecond, intensities ($J/cm^2 \cdot sec$) can be achieved that are high enough for occurrence of non-linear interactions between the beam and the transmitting material. This may lead to new damage mechanisms. Experimental data for the ultrashort, sub-picosecond regime are only beginning to become available. {3,4}

In this report, we are investigating how to determine specific molecular species and the biochemical reactions that lead to cellular damage. In order to try to understand the damage on a biochemical level, we have looked at the best characterized regime of pulse lengths which is for durations greater than one

microsecond. In this regime, the damage is almost certainly due to temperature rises of the cellular material. This allows us to use rate kinetics to describe ways for analyzing data to get possible reaction mechanisms.

This report first briefly reviews Henriques' original work on cellular damage caused by temperature rises. The next section describes inconsistencies in using Henriques' simple model and how to correct them. Finally, possible reaction pathways are investigated for their kinetic behavior and types of experiments are described that could lead to kinetic information that would be valuable for formulating actual reaction mechanisms.

II. Review of Henriques One-Step Thermal Damage Model

The Arrhenius Damage Integral, introduced in the context of thermal injury to biological tissue by Henriques {5}, assumes that the damage mechanism follows a standard reaction rate expression:

$$\frac{d\Omega}{dt} = Pe^{-\Delta E/R(T_o+\Delta T)} \qquad (1)$$

where $d\Omega/dt$ is the rate at which damage occurs. This expression can be useful for determining the quantities ΔE and P which are related to the thermodynamics of the underlying process through the following approach. Equation (1) can be rewritten as:

$$\Omega = \int Pe^{-\Delta E/R(T_o+\Delta T)}\,dt \qquad (2)$$

ΔE and P can be determined if the time dependence of ΔT is known for a laser pulse that is known to cause damage. Under these circumstances, it is customary to set $\Omega=1$, representing that the system has gone from an undamaged state to a damaged state. As a simple example, if the laser pulse causes a temperature rise ΔT that remains constant for a time Δt that is long compared to its rise or fall times, then the integral is easily evaluated

$$\Omega = 1 = Pe^{-\Delta E/R(T_o+\Delta T)}\Delta t \qquad (3)$$

If this can be done with different pulses that result in different, but known ΔT and Δt, then ΔE and P can be determined.

As shown by Henriques {5}, using standard reaction rate theory, ΔE is the activation energy of the process and the pre-exponential P is related to the entropy change, ΔS

$$P = \frac{k_B T}{h} e^{\Delta S/R} \qquad (4)$$

ΔS can be obtained from P through

$$\Delta S = R \ln(P \times h/k_B T) \qquad (5)$$

Using R=1.987 cal/mol•°K, $h=6.6\times10^{-34}$J•s, $k_B=1.4\times10^{-23}$J/°K, and T≈350 K, gives

$$\Delta S = 1.987 \text{ cal/mol}\cdot{}^\circ\text{K} \ln(P \times 1.47\times10^{-13} \text{ sec})$$

$$= 4.57 \text{ cal/mol}\cdot{}^\circ\text{K} \log(P \times 1.47\times10^{-13} \text{ sec}) \qquad (5a)$$

This model has been used by a variety of researchers. The results of some of these studies are the following:

Henriques {5}: $P = 3.1 \times 10^{98}$ sec^{-1} → ΔS = 391 cal/mol•°K

ΔE = 150 Kcal/mol

(porcine skin burns from non-laser, direct surface heating with exposures from 1 second to 25,000 seconds)

Welch and Polhamus {6}: $P = 1.3 \times 10^{99}$ sec^{-1} → ΔS = 394 cal/mol•°K

ΔE = 150 Kcal/mol

Takata {7}:

For $T \leq 323$ K $\quad P = 4.322 \times 10^{64}\ \text{sec}^{-1} \rightarrow \Delta S = 240$ cal/mol•°K

$\Delta E = 99$ Kcal/mol

For $T > 323$ K $\quad P = 9.389 \times 10^{104}\ \text{sec}^{-1} \rightarrow \Delta S = 420$ cal/mol•°K

$\Delta E = 158$ Kcal/mol

(epidermal skin burns created by HeNe laser, λ=632.8 nm)

Birngruber et. al. {8}:

$P = 3 \times 10^{44}\ \text{sec}^{-1} \rightarrow \Delta S = 137$ cal/mol•°K

$\Delta E = 70$ Kcal/mole

(from the work of Vassiliadis et. al. for retinal injury with pulse lengths of 1 msec to 300 msec)

It can be seen from the values of ΔE and ΔS above that there are significant differences in the values obtained, though they are all of the same order of magnitude.

The parameters ΔE and ΔS have precise meanings in chemical reactions in terms of the activation energy for a reaction leading directly to the observed product. This assumes knowledge of a well defined chemical reaction pathway and it is not clear that this is straightforwardly applicable for the complicated processes involved in laser induced retinal damage. The work reports on how the details of a reaction pathway might be discerned from data such as that described above.

In addition, Eqs. {1-3) imply that even at normal temperatures T_o, damage should be occurring, i.e. $d\Omega/dt > 0$. Yet buildup of damage does not occur at normal temperatures. Thus, simple use of Eq. (1) ignores the kinetics of a repair process.

Given the problems just discussed, there are nevertheless reasons to maintain the Arrhenius Damage Integral formalism. Most importantly, the underlying mechanism causing the damage is some type of biological process

following statistical thermodynamics and exponential dependencies on temperature can be expected, though possibly in a more complicated fashion than expressed in Eq. (1). Also, there is a curious coincidence that when an expression as simple as Eq. (1) is used to fit the data, the value of ΔE is in the range of 50-150 Kcal/mole which is appropriate for energies for protein denaturation, a very possible cause for the observed damage. Thus, it is worthwhile to develop a theoretical model that is a more realistic version of Eq. (1) and which will incorporate features to correct some of the faults pointed out above.

III. Inclusion of A Repair Process

The Arrhenius Integral used by Henriques is the accepted way for analyzing chemical reaction and is used to gain information on the underlying mechanisms. Henriques assumed that thermal damage occurs by a zero order change in a molecule and we now expand on his analysis. The simplest extension is to include a repair, or back reaction. To attempt to relate the model to actual biochemical molecules and reactions, we change the notation. The concentration of molecules in the undamaged state will be denoted by C_U, C_D will be the concentration of molecules in the damaged state, and C_0 is the total concentration of molecules. In a first order process (Henriques' is zero-order)

$$\frac{dC_U}{dt} = -K(T;t)\,C_U \qquad (6)$$

where K(T;t) is the first order rate constant which depends explicitly on temperature, and which has a time dependence because of the laser induced heating. From reaction rate theory {9}

$$K = \frac{k_B T}{h} e^{\Delta S/R} e^{-\Delta H/RT} \qquad (7)$$

The connection between K, C_U, and Ω is then

$$\Omega(t,T) = -\ln\left(\frac{C_A(t)}{C_o}\right) = {}_o\int^t K dt' \qquad (8)$$

When the process starts, $C_U(t=0)=C_o$ and therefore $\Omega(t=0)=0$. When $C_U/C_o=36.8\%$, then $\Omega=1$. Therefore, the observation of a damage endpoint is interpreted as signifying that 63.2% of some type of molecule has undergone a first order process to a damaged state.

Thus, the Arrhenius Damage Integral is used with experimental data to obtain quantities that are related to the thermodynamics of the underlying biochemical reactions. However, the full value of the Arrhenius Damage Integral may not be currently realized because of the assumption that the damage process can be described by a one step reaction. As Henriques stated {5, p.500}, "With no intention whatsoever of inferring that the thermal effects on living protoplasm can be ascribed to the alteration of any single protein, it is of value to make for the moment this extreme oversimplification..."

The aim of this research is to extend the model of Henriques for damage to biological tissue caused by temperature increases. Henriques' model includes only a single, thermally activated process that leads from undamaged to damaged tissue. However, unlike the stable condition found in healthy tissue, a single damage causing step does not allow for an equilibrium situation but instead implies that even at normal body temperature there should be a slow but continuous build up of damage. The simplest extension is to add a repair process. The two state model involving forward (damaging) and back (repair) unimolecular reactions

$$C_U \underset{k_b}{\overset{k_f}{\rightleftharpoons}} C_D$$

is described by the two kinetic equations

$$\frac{dC_U}{dt} = -k_f C_U + k_b C_D \qquad \frac{dC_D}{dt} = k_f C_U - k_b C_D \qquad (9)$$

It is assumed that the molecules are initially in the undamaged state so that

$C_U(t)+C_D(t)=C_o \equiv C_U(t=0)$. With this assumption, these equations can be solved analytically and give the time dependence of the concentrations

$$C_U(t)=C_U(0)\,\frac{k_b+k_f e^{-(k_f+k_b)t}}{k_f+k_b} \tag{10a}$$

and

$$C_D(t)=C-C_U=C_U(0)\,\frac{k_f-k_f e^{-(k_f+k_b)t}}{k_f+k_b} \tag{10b}$$

In equilibrium, C_U and C_D are independent of time and the ratio of the concentrations ξ is

$$\xi^{Eq}(T)=\frac{C_D}{C_U}=\frac{k_f}{k_b} \tag{11}$$

Equation (11) shows that if the rate constants have a temperature dependence then so too will the equilibrium constant $\xi^{Eq}=\xi^{Eq}(T)$.

The physical justification for the existence of the back reaction, the repair process is that proteins cease to function when their three dimensional native state conformation is only moderately disrupted. This requires only slight unfolding of their complex shapes and occurs at temperature increases of less than 50 K. For example, the unfolding or "melting temperature" of the commonly studied protein bovine pancreatic ribonuclease A (RNase A) is 62°C, only 25°C above body temperature {10}. This is the same range of temperature at which threshold levels for thermal damage occur and these temperature increases are not high enough to break covalent bonds in proteins, but only weaker interactions like hydrogen bonds. The primary sequence of amino acids is still intact and this contains all the information necessary for a protein to repair itself by refolding to its functioning native state. Unfolding raises the entropy of the molecule which makes unfolding more likely as the temperature increases. Thus, k_f is expected to increase with temperature and k_b is expected to decrease as the temperature rises, shifting the equilibrium towards a greater fraction for C_D. We can use Henriques' approach to ascertain information about the thermodynamic

parameters of the rate constants for the damage and the repair process.

In general, rate constants depend on several factors

$$k = A\nu e^{-\Delta G^{\ddagger}/RT} \tag{12}$$

The parameter ν is the barrier attempt frequency for a molecule. It is due to thermal vibrations of reactive normal modes that can lead to occurrence of the process or reaction. At body temperature, biomolecular normal modes have $\nu \approx K_B T/h \approx 10^{13} s^{-1}$. The factor $\exp(-\Delta G^{\ddagger}/RT)$ is the probability per attempt that the molecular system has the necessary free energy to get to the top of the barrier. The parameter A is the probability that the system will go to the new state on the other side of the barrier if it gets to the top. In an adiabatic process in which the initial and final states of the system are strongly coupled by many degrees of freedom, as is the case here for the denaturing of proteins, $A \approx 1$ and will be left out of the rest of this report.

In equilibrium, substituting Eq. (12) into Eq. (11) gives

$$\xi^{Eq}(T) = \frac{C_D}{C_U} = \frac{k_f}{k_b} = \frac{\nu_f e^{-\Delta G_f^{\ddagger}/RT}}{\nu_b e^{-\Delta G_b^{\ddagger}/RT}} \tag{13}$$

Since all normal modes ν have the same value, the equilibrium constant then reduces to the expected Boltzmann factor

$$\xi^{Eq}(T) = \frac{C_D}{C_U} = e^{-\Delta G_o^{\ddagger}/RT} \tag{14}$$

where $\Delta G_o^{\ddagger} = \Delta G_f^{\ddagger} - \Delta G_b^{\ddagger} = \Delta G_D^{\ddagger} - \Delta G_U^{\ddagger}$. This expression allows us to focus on the thermodynamic parameters of the process and permits an investigation of how the temperature dependence of the equilibrium is related to the thermodynamics of the underlying molecular process. We use $\Delta G_o^{\ddagger} = \Delta H_o^{\ddagger} - T\Delta S_o^{\ddagger}$ where $\Delta H_o^{\ddagger} = \Delta H_D^{\ddagger} - \Delta H_U^{\ddagger}$ and $\Delta S_o^{\ddagger} = \Delta S_D^{\ddagger} - \Delta S_U^{\ddagger}$.

Biomolecules such as proteins, are marginally stable at body temperature.

Though both $\Delta H_0^{\ddagger}$ and $\Delta S_0^{\ddagger}$ are significant, they are nearly balanced to give a small $\Delta G_0^{\ddagger}$. To get an idea of the delicate balance, a typical hydrogen bond is approximately 20 kJ/mol and there are dozens contributing to the stabilization of a protein. However, the free energy required to denature proteins is approximately 0.4 kJ/mol of amino acid residues {11} so that a protein consisting of 100 residues is stable by only 40 kJ/mol, which is much smaller than contributions from hydrogen bonds, or many other interactions. For barnase, the RNAse from B. *amyloliquefaciens*, which has 110 residues, $\Delta G_0 \approx 44$ kJ/mol.{12} For proteins, both ΔH_0 and ΔS_0 can be temperature dependent {13}

$$\Delta H_o(T) = \Delta H_o(T_o) + \int_{T_o}^{T} \Delta C_p dT \qquad (15a)$$

$$\Delta S_o(T) = \Delta S_o(T_o) + \int_{T_o}^{T} \frac{\Delta C_p}{T} dT \qquad (15b)$$

where $\Delta H_0(T)$ and $\Delta S_0(T)$ are the differences in the enthalpy and the entropy in going from the undamaged native state to the damaged denatured state at the temperature T, T_0 is a reference temperature, and $\Delta C_p = C_p(D) - C_p(U)$. In the temperature range for threshold levels for thermal damage, which is the same as for protein denaturation studies, 37°C → 80°C, ΔC_p has been found to be nearly constant {14} so equations 15a and 15b reduce to

$$\Delta H_o(T) = \Delta H_o(T_o) + \Delta C_p(T - T_o) \qquad (16a)$$

$$\Delta S_o(T) = \Delta S_o(T_o) + \Delta C_p \ln\left(\frac{T}{T_o}\right) \qquad (16b)$$

At room temperature, T_0=25°C, ΔH_0 and ΔS_0 are near zero but ΔC_p has a large value of 4-8 kJ/mol·K {13}. As the temperature increases above 25°C, ΔH_0 and ΔS_0 increase rapidly, but continue to compensate. Using ΔC_p=6 kJ/mol, at body temperature of 37°C, ΔH_0(37°C)≈72 kJ/mol and $T\Delta S_0$ is approximately the same,

showing the marginal stability of proteins and how thermal denaturation might be responsible for threshold level damage from thermal effects.

Henriques model involved only a single thermally activated step causing damage, Ω. His expression for the rate at which damage occurs, $d\Omega/dt$ is (ΔE in Henriques' paper is now substituted for by ΔH_U)

$$\frac{d\Omega}{dt} = Pe^{-\Delta H_U/RT} \tag{17}$$

Henriques then rewrote Eq. (17) as an integral expression over time

$$\Omega = \int Pe^{-\Delta H_U/RT} dt \tag{18}$$

As shown in section II describing Henriques' model, ΔH_U and P can be determined if the time dependence of ΔT is known for a laser pulse that is known to cause damage and values are given for ΔS and ΔH (ΔE) in that section.

In order to investigate Henriques' one step damage process in terms of the formal equations of kinetic theory of Eqs. (9, 10, and 12), we rewrite Eq. (17)

$$\frac{d\Omega}{dt} = \nu e^{-\Delta G_U/RT} = k_f \tag{19}$$

The integral expression then becomes

$$\Omega = \int \nu e^{-\Delta G_U/RT} dt = \int k_f dt \tag{20}$$

We now have expressions for the build up of damage Ω. In order to make these expressions, Eq. (18) or Eq. (20), relevant to experimental measurements the parameter Ω is chosen to be equal to 1.0 when damage is experimentally observable. In addition, to relate the Henriques model to some kind of thermodynamic system, we assume for now that $\Omega = C_D = C_o - C_U$ which gives $d\Omega/dt = -dC_U/dt$. This allows us to rewrite Eq. (17) or Eq. (19) as a first order one step reaction

$$\frac{dC_U}{dt} = -k_f(T; t)\, C_U \tag{21}$$

Since ${}_0\!\int^t dC_U/C_U = \ln\,[C_U(t)/C_U(0)] = -{}_0\!\int^t k_f dt = -\Omega$, this implies that at damage, when Ω is designated to be 1.0, that $C_U(t)/C_U(0) = .368$. Inserting $C_o=C_U(0)$ gives $C_D(t)=C_o-C_U(t)=.63$ and when damage is observed $\xi(T;t)=C_D(t)/C_U(t)=1.7$.

The results of this analysis for ΔH_f, ΔS_f and ξ can also be used to determine if they are consistent with the lack of observable damage at normal body temperature over the lifetime of an individual. In order to evaluate $C_U(t)$ and $C_D(t)$ at another temperature, the temperature dependence of k_f due to the temperature dependence of ΔG_U must be known. The simplest approach is to assume that both ΔH_U and ΔS_U are temperature independent and therefore the temperature dependence of ΔG_U is purely due to its explicit T dependence.

We have just shown the plausibility of thermal denaturation as the mechanism for threshold thermal damage. Thermal denaturation is reversible and many proteins refold (repair) as the temperature is lowered back to body temperature. Introduction of the repair process in Eq. (9) is therefore justified on physical grounds and also leads to the correct temperature dependence for increasing damage as the temperature rises. Additionally, the solutions to the kinetic equations, Eq. (9) and (10), are straightforward. Unfortunately, the addition of a repair process alone does not explain how the kinetics of the molecular processes lead to the observed cellular damage. The discrepancy arises from the difference in time scales. The time scale for the elevated temperature for pulses longer than a microsecond is approximately that of the laser pulse. The time scale for large scale protein molecular rearrangement, folding or unfolding, is on the order of milliseconds to seconds and therefore can occur during the time that the temperature is elevated. However, minimal visible lesions (MVL) can take minutes or hours to appear. This is long after the temperature has returned to normal and any proteins that are going to refold will have done so. Therefore, there must be additional steps leading from damaged molecules to observable cellular damage. There are two possible explanations:

1)A critical fraction of molecules do not get repaired. The effect of this is additional damage to other components of the cell which slowly builds up and causes observable damage. This possibility can be distinguished by looking for experimental markers (e.g. circular dichroism markers of protein denaturation) which appear within a second, or less, after the end of the pulse, and do not go away as MVL forms.

2)Most molecules that are directly damaged by the laser pulse do get repaired within seconds, but during the few seconds that they are disabled they affect other molecules in the cell, which in turn lead to MVL. This possibility can be investigated by determining if the immediate experimental markers of damage, such as CD, return to normal in the following seconds while the MVL builds on a slower time scale.

In order to further investigate these possibilities and relate them to thermodynamic quantities as in Henriques' original model, further analysis of reaction mechanisms is necessary.

IV. Rate Kinetics

In order to attempt to map out molecular pathways leading from temperature rises to cellular damage we investigate how reaction kinetics may be applied to multi-step biochemical reactions. Henriques one step damage process of Eq. (1) is of the form of a zero order reaction

$$\frac{dc}{dt}=k \tag{22}$$

A reaction of this form has a rate that is independent of the concentration of reactants. This kind of reaction cannot be explained by a mechanism that involves a single molecule experiencing damage unless there is another process that introduces new undamaged molecules at the same rate so that the concentration of undamaged molecules remain the same. An example of a biochemical reaction that exhibits this zero-order behavior is the conversion of ethanol to acetaldehyde by the oxidizing agent nicotinamide adenine dinucleotide

(NAD^+) in the presence of the liver enzyme alcohol dehydrogenase (LADH):

$$CH_3CH_2OH + NAD^+ \xrightarrow{\text{LADH}} CH_3CHO + NADH + H^+$$

If there is an excess of alcohol compared to the enzyme and if the NAD^+ is buffered via metabolic reactions that rapidly restore it, the rate of the reaction in the liver is zero-order

$$\frac{d[CH_3CHO]}{dt} = -\frac{d[CH_3CH_2OH]}{dt} = k \qquad (23)$$

This reaction is zero-order because the rate is limited by the concentration of the enzyme which is neither a reactant nor a product, and because the concentration of NAD^+ is continually restored.

A zero-order reaction mechanism could conceivably explain thermal damage to retinal cells if the damage mechanism is catalyzed by an agent that is created by the temperature rise. The actual molecule that causes the observable cellular damage is not directly affected by the temperature rise but changes its concentration due to the action of the catalyzing agent which is activated by the temperature rise. A possible mechanism could be the disrupting of the membrane that encloses melanosomes. Melanosomes contain melanin that is packaged with protein coats that restrict the direct interaction of the melanin with the surrounding cellular medium. There is evidence that melanin can adversely affect cellular components if the membrane that encloses the melanosomes is disrupted. If the exposed melanin is photo-excited with a laser pulse that pumps the free radical activity of the melanosomes, significant oxidation of NAPDH was found in an *in vitro* solution.{15} The same mechanism in a retinal cell could lead to damage as undamaged molecules in the cell randomly make contact with the small area of exposed melanin. In this scenario, the temperature dependence of the rate constant in the Henriques expression of Eq. (1) would be due to the temperature dependence of the integrity of the surrounding membrane. Interestingly, the values for ΔH and ΔS are in the general range for protein denaturation as discussed earlier and therefore the extremely simple zero-order

process might underlie the damage process. Furthermore, this zero-order process allows the same exposed sites on melanosomes to cause damage on many molecules in the cell which fits within the general idea of "biological amplification".

The plots in Figs. (2-8) show the temperature on a melanosome surface using a granular model. The computer code uses a model {16} with the following properties. The laser beam has a gaussian spatial profile transverse to the direction of propagation with a $1/e^2$ radius of 11 μm. It is incident on a material with the thermal properties of water. The medium is effectively infinite in the x-y directions transverse to the beam and is one RPE cell in thickness, 15 μm. The laser energy is absorbed by melanosomes that are modelled as 1 μm radius spheres with an absorption coefficient of 2000 cm^{-1}. The melanosomes are scattered randomly with a number density of $100/600\mu m^3$ where 600 μm^3 ($20\mu m \times 20\mu m \times 15\mu m$) is the volume of one RPE cell. The temperature is plotted {17} for the surface of a melanosome that is located at a distance from the beam center such that it experiences a beam intensity that is 1/2 that of the intensity at the beam center. The temperature rise of the melanosome depends not only on the energy it absorbs from the laser but also the energy absorbed by other melanosomes nearby. For each figure, the ED_{50} fluence for that pulse duration is used. Using the Henriques damage integral of Eq. (2) with values from Welch and Polhamus {6} of $P=1.3 \times 10^{99}$ sec^{-1} and ΔE=150 Kcal/mole, we have shown {16} that Eq. (2) fits the ED_{50} data fairly well, i.e. the values of the fluence needed to get $\Omega=1$ are close to those that were experimentally determined. However, this was done by calculating the temperature rises outside the melanosome in the cellular medium. It would be interesting to use these values of P and ΔE along with the temperature rise on the surface of the melanosome in Eq. (2) to compute a new ED_{50} to compare with those plotted in Figure 1. Conversely, it would also be useful to determine the P and ΔE that would lead to good agreement, i.e. use the ED_{50}'s from the experimental data and determine the P and ΔE that will give a value of one for Ω in Eq. (2). Also, using the

temperature on the surface of the melanosome may affect the curvature of the plot of fluence vs. laser pulse duration in a way that can either help or hurt the fit to the experimental ED_{50}'s.

Another step in determining if disruption of melanosome membranes may be important for thermal damage is to experiment on isolated melanosomes to determine if the membranes disrupt when subjected to the temperature rises expected at the surface of the melanosomes from ED_{50} fluence. The experiment would involve isolating melanosomes and placing them in water, then warming the mixture and assaying for membrane disruption. Additional experiments would involve placing intact melanosomes in a solution such as the NADPH used by Glickman et. al. and determining if the melanosomes become photo-reactive with the NADPH as they are heated. Other types of biochemical molecules should also be tested to see if they are affected by heated melanosomes.

Though the zero-order, melanosome membrane disruption mechanism may be the actual mechanism for damage, we now investigate other reaction mechanisms. The next simplest reaction kinetics is first order

$$A \xrightarrow{k_1} B$$

which is described by the kinetic equations

$$-\frac{d[A]}{dt} = \frac{d[B]}{dt} = k_1 [A] \tag{24}$$

This is the reaction mechanism of Eq. (6) in which $[A]=C_U$ and $[B]=C_D$. As shown earlier, the buildup of damage follows Eq. (8). The integration can be performed simply if k is constant. Since k is expected to be temperature dependent, as the temperature changes so will k. Since temperature rise and fall happen on microsecond time scales, for pulse lengths longer 10 μseconds the temperature is effectively a step function in time, as used in evaluating the integral in Eq. (3), and we get

$$C_U = C_o e^{-k_1 t} \quad \rightarrow \quad C_D = C_o (1 - e^{-k_1 t}) \tag{25}$$

In the presence of a back (repair) mechanism, the reaction kinetics and time dependence are those given above in Eqs. (10).

The next level of complexity is second-order reactions in which the concentration of two reactant molecules determines the reaction rate. Second order reaction fall under two broad categories

I. A + A -----> P

II. A + B -----> P

For category I the reaction kinetics are given by

$$\frac{d[A]}{dt} = -k_2 [A]^2 \tag{26}$$

If k_2 is independent of time, the integrated form of this second-order rate equation is

$$\frac{1}{[A]} - \frac{1}{[A]_o} = k_2 t \tag{27}$$

where $[A]=[A(t)]=[A]_o-x$ and x is the concentration of A which has reacted. For category II the reaction kinetics are given by

$$\frac{d[P]}{dt} = -\frac{d[A]}{dt} = -\frac{d[B]}{dt} = k_2 [A]\,[B] \tag{28}$$

With a change of variables this can be integrated to

$$\frac{1}{[A]_o - [B]_o} \ln \frac{[B]_o [A]}{[A]_o [B]} = k_2 t \tag{29}$$

with the concentration of product at any time given by $[P]=[A]_o-[A]=[B]_o-[B]$. Though not expected for molecules in a cell, if the initial concentrations of the

different species A and B are the same, $[A]_0=[B]_0$, then the kinetics behave the same as for a Type I second order reaction and give the same expressions for the build-up of product as a function of time. For second order reactions, the rate of build-up of product is seen to depend not only on the rate constant but also on the initial concentrations.

There are reactions that are third order, and higher, but we now switch to investigating complex reactions with more than one step. The vast majority of biochemical reactions are complex. We will look specifically at a reaction process that involves a series of steps and examine how the concentrations of the various components change with time. We also look at the limiting cases where one of the steps is slow enough to be rate limiting and the danger of then missing the other, faster step because of the difficulty of measuring the quickly changing concentrations that it entails.

For a series reaction pathway, the molecules that are initially perturbed by the temperature rise cause damage by reacting with other species. This can be represented by

$$A \xrightarrow{k_1} B \xrightarrow{k_2} C$$

If the reactions are first order, then

$$\frac{d[A]}{dt} = -k_1 [A] \quad \rightarrow \quad [A] = [A]_o e^{-k_1 t} \tag{30}$$

and

$$\frac{d[B]}{dt} = k_1 [A] - k_2 [B] = k_1 [A]_o e^{-k_1 t} - k_2 [B] \tag{31}$$

The solution of this differential equation, with the initial condition of $[B]_0=0$ is

$$[B] = \frac{k_1 [A]_o}{k_2 - k_1} [e^{-k_1 t} - e^{-k_2 t}] \tag{32}$$

Similarly,

$$\frac{d[C]}{dt} = k_2[B] = \frac{k_1 k_2 [A]_o}{k_2 - k_1}[e^{-k_1 t} - e^{-k_2 t}] \tag{33}$$

This can be solved by separation of variables, and with the initial condition $[C]_o=0$, the build up of concentration of [C] is given by

$$[C] = [A]_o \left\{1 - \frac{1}{k_2 - k_1}[k_2 e^{-k_1 t} - k_1 e^{-k_2 t}]\right\} \tag{34}$$

The kinetic curves for reactant [A], intermediate [B], and product [C] for this series reaction have the general form of Fig. 9 when the rate constants are of similar size $k_1 \approx k_2$. The curve of [C] versus t must start out with zero slope since initially $[B]_o=0$. Asymptotically, [C] approaches a constant value and therefore the curve of [C] must have an inflection point, which occurs when [B] is at its maximum. The specific time for this inflection point of [C] relative to the disappearance of [A] and build up of [C] depend carefully on the relative sizes of k_1 and k_2. For the extreme condition $k_1 \gg k_2$, the concentrations vary with time as in Fig. 10 and for the opposite extreme of $k_2 \gg k_1$, the curves look like Fig. 11. As can be seen, the slowest step determines the time for the build up of the product C. These dramatically different time courses allow the distinguishing of a series reaction from a one step reaction most easily if $k_1 \approx k_2$. Unfortunately, if $k_1 \gg k_2$, the concentration of the first reactant [A] disappears very quickly with a time constant of $1/k_1$. Concentration measurements of the longer lived [B] and [C] could lead to a model of a single step, first-order process and the importance of the initial formation of [A] directly from the laser pulse could be overlooked. This would be unfortunate in that A might be a molecule whose concentration or reaction to laser energy can be controlled by biochemical, pharmaceutical, population selection, hypothermal preparation, or other means to lessen (or increase) the damage from laser exposure. Likewise, in the case for which $k_2 \gg k_1$, the concentration of the intermediate B never rises

to an appreciable level compared to A and C. Again, an important step in the damage process can be overlooked. To avoid the possibility of mistaking a multi-step series process for a single step process, careful measurements must be made that look for small amounts and small changes in the concentrations of many different species of molecules. Even when some of the reacting species are identified, additional molecules and steps may be involved that are harder to discern but could be of equal importance.

V. Experimental Techniques

In order to discern the various reaction kinetics, e.g. zero-order, first-order, series, etc., experimental measurements must be made of concentration dependencies as a function of time. The order of a reaction can then be determined by plotting as a function of time the changing concentration, the log of the concentration, the inverse of the concentration, etc. and picking out which appears closest to a straight line. Additional methods for determining the order of a reaction involve differentiation of the data (velocity), method of initial rates, changing the initial concentrations, and supplying reagents in excess.

The actual measurements of the concentrations can done with a variety of methods. The experimental technique used to measure the increased reactivity of melanosomes that have their membranes disrupted are discussed in Ref. {15}. There are a variety of standard techniques that can be used, initially *in vitro* in order to isolate the critical molecules and biochemical reactions.

VI. Conclusion

Various reaction kinetics are described above. Interestingly, it appears that the simple zero-order kinetics may be the best explanation. The Henriques damage integral does an excellent job of matching the experimental ED_{50}'s in the time range from 10^{-4} to 1 seconds. The proposed mechanism of membrane disruption of melanosomes must be investigated further to see if it is responsible for

catalyzing damaging molecular changes. It is known that when melanin is irradiated at levels below the thermal damage level {18}, the melanin radicals are excited to a state that will oxidize biochemical substrates such as ascorbic acid {19,20}. Since this happens at irradiation levels below thermal damage levels, it is consistent with the thermal damage levels being determined by the requirement of disrupting the membrane, as discussed above. If this is the underlying mechanism for threshold levels of thermal damage, it would be a direct example of biological amplification in which the creation of small scale changes on micron size melanosomes catalyzes reactions that eventually lead to observable cellular death.

References

1)C. P. Lin and M. W. Kelly, Ultrafast time-resolved imaging of stress transient and cavitation from short pulsed laser irradiated melanin particles, in *Laser-Tissue Interaction VI*, S. L. Jacques, Editor, Proc. SPIE 2391, 294-299, 1995.

2)B. S. Gerstman, C. R. Thompson, S. L. Jacques, and M. E. Rogers, Laser Induced Bubble Formation in the Retina, Lasers in Surgery and Medicine, **18**(1), 10-21, 1996.

3)C. P. Cain, C. D. DiCarlo, P. K. Kennedy, D. D. Noojin, R. E. Amnotte, and W. P. Roach, *In vivo* laser-induced breakdown in the rabbit eye, in *Laser-Tissue Interaction VI*, S. L. Jacques, Editor, Proc. SPIE 2391, 41-47, 1995.

4)D. X. Hammer, R. J. Thomas, G. D. Noojin, B. A. Rockwell, and A. Vogel, Ultrashort pulse laser induced bubble creation thresholds in ocular media, in *Laser-Tissue Interaction VI*, S. L. Jacques, Editor, Proc. SPIE 2391, 30-40, 1995.

5)F. C. Henriques, Studies of Thermal Injury, Archives of Pathology, **43**, 489-502, 1947.

6)A. J. Welch and G. D. Polhamus, Measurement and Prediction of Thermal Injury on the Retina of the Rhesus Monkey, IEEE Trans. Biomed. Eng. **BME-31 (10)**, 1471-1481, 1984.

7)A. N. Takata, Development of Criterion for Skin Burns, Aerospace Medicine, 45, 634-637, 1974.

8)R. Birngruber, F. Hillenkamp, and V. P. Gabel, Theoretical Investigations of Laser Thermal Retinal Injury, Health Physics, 48, 781-796, 1985.

9)I. Tinococ, Jr., K. Sauer, and J. C. Wang, Physical Chemistry, Third Edition, Prentice Hall, New Jersey, 1995.

10)P. H. von Hippel and K. Y. Hong, J. Biol. Chem. **10**, 3911, 1965.

11)D. Voet and J. G. Voet, Biochemistry, John Wiley & Sons, New York, p. 175, 1990.

12)J. T. Kellis Jr., K. Nyberg, D. Dali, A. R. Fersht, Contribution of hydrophobic interactions to protein stability, Nature **333**, 784-786, 1988.

13)T. Alber, Mutational effects on protein stability, Ann. Rev. Biochem., **58**, 765-798, 1989.

14)P. L. Privalov and S. J. Gill, Adv. Protein. Chem., **39**, 191-234, 1988.

15)R. D. Glickman, S. L. Jacques, J. A. Schwartz, T. Rodriguez, k. Lam, and G. Buhr, Photodisruption increases the free radical reactivity of melanosomes isolated from retinal pigment epithelium, SPIE (International Society for Optical Engineering) Proceedings, 2681(62), 1996.

16)C. R. Thompson, B. S. Gerstman, S. L. Jacques, and M. E. Rogers, Melanin Granule Model for Laser Induced Thermal Damage in the Retina, Accepted for publication in Bulletin of Mathematical Biology, 1996.

17)Courtesy of Dr. C. R. Thompson.

18)American National Standard for the Safe Use of Lasers, ANSI **Z.136.2-1993**, New York: Am. Nat. Std. Inst., 1993.

19)R. D. Glickman and K. -W. Lam, Oxidation of ascorbic acid as an indicator of photooxidative stress in the eye, Photochem. Photobiol. **55**, 191-196, 1991.

20)K. -W. Lam and R. D. Glickman, Prevention of light-induced free radical production from melanin granules by ascorbic acid, in: K. Yagi, E. Niki, T. Yoskikawa, eds., Oxygen Radicals, 633-636, Elsevier Science Publishers, Amsterdam, 1992.

Figures

Figure 1) Experimental ED_{50} threshold corneal fluence for retinal damage as a function of laser pulse duration. As a result of beam focusing, the fluence is increased by approximately 10^5 when it reaches the retina. Many different wavelengths in the visible and near-IR are included, accounting for much of the spread.

Figures 2-8) Plots of temperature vs. time at the surface of a melanosome. Each figure is for a different laser pulse duration. All figures have the following:

1)Beam radius ($1/e^2$) = 11 μm.

2)Location of melanosome such that fluence is ½ that at beam center.

3)Melanosome properties: radius = 1 μm, absorption α = 2000 cm^{-1}.

4)Melanosome number density = $100/600\mu m^3$ (=100 per RPE cell).

Figure #	Pulse Duration (sec)	Retinal ED_{50} Fluence (J/cm^2)
2	10^{-6}	1
3	10^{-5}	1
4	10^{-4}	2
5	10^{-3}	9
6	10^{-2}	100
7	0.1	500
8	1.0	1500

Figures 9 -11) Plot of concentrations of [A], [B], and [C] during the time of the reaction for the series reaction of Eqs. (30) - (34):

Fig. (9): $k_1 \approx k_2$

Fig. (10): $k_1 \gg k_2$

Fig. (11): $k_1 \ll k_2$

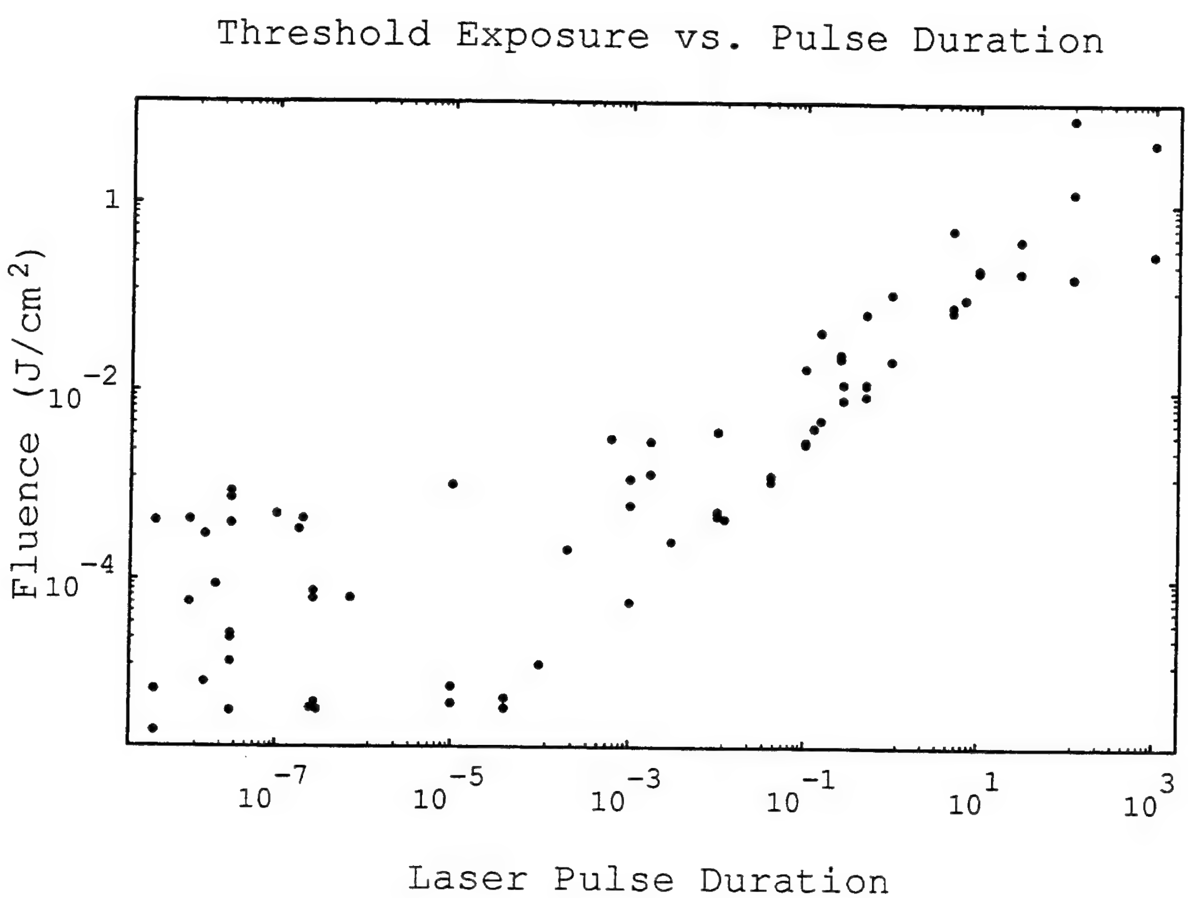

Threshold Exposure vs. Pulse Duration
1
10^{-2}
10^{-4}
Fluence (J/cm^2)
10^{-7}
10^{-5}
10^{-3}
10^{-1}
10^{1}
10^{3}
Laser Pulse Duration

Figure 1

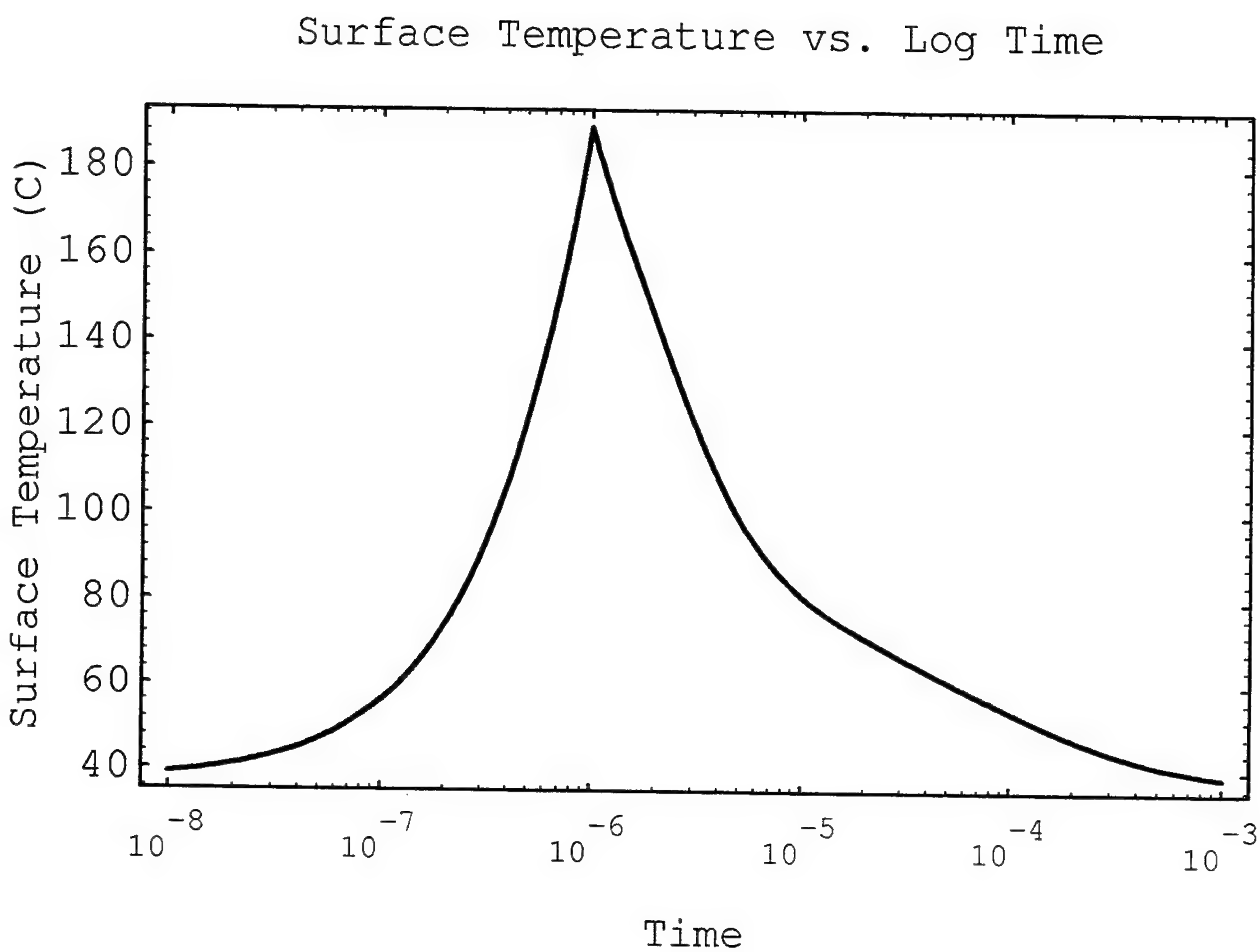

Surface Temperature vs. Log Time
Surface Temperature (C)
180
160
140
120
100
80
60
40
10^-8
10^-7
10^-6
10^-5
10^-4
10^-3
Time

Figure 2

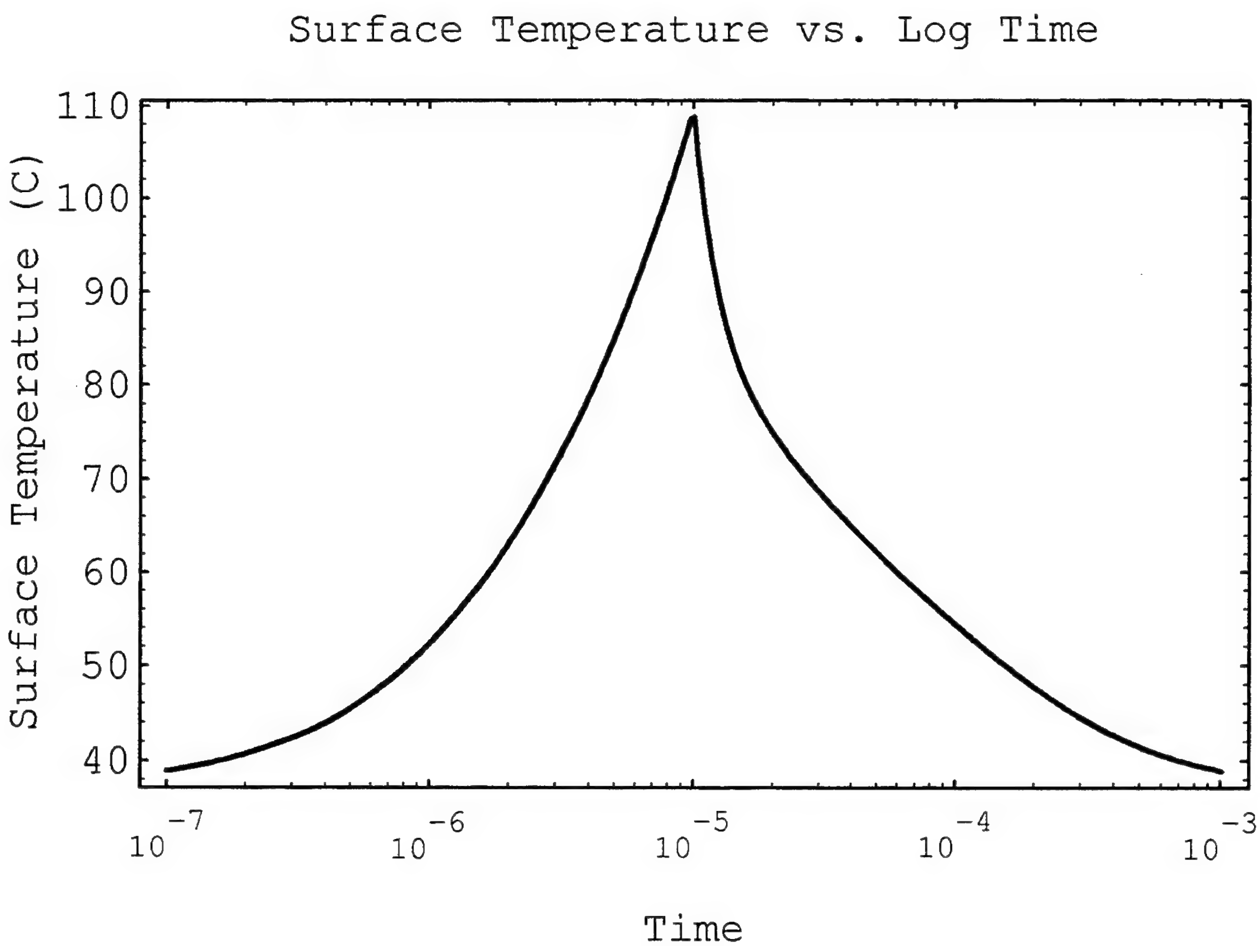

Surface Temperature vs. Log Time
Surface Temperature (C)
110
100
90
80
70
60
50
40
10 −7
10 −6
10 −5
10 −4
10 −3
Time

Figure 3

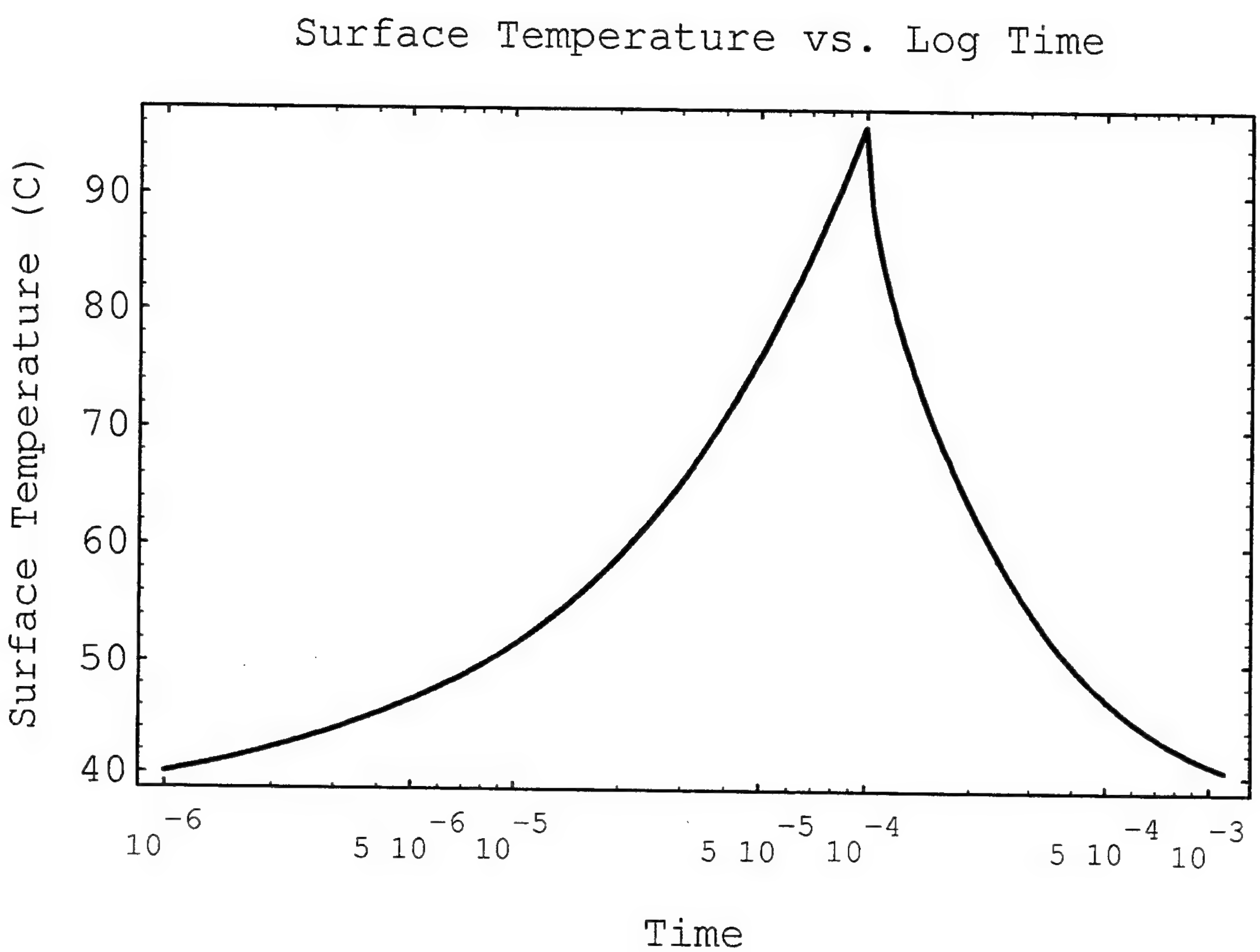

Surface Temperature vs. Log Time
Surface Temperature (C)
90
80
70
60
50
40
10 -6
5 10 -6
10 -5
5 10 -5
10 -4
5 10 -4
10 -3
Time

Figure 4

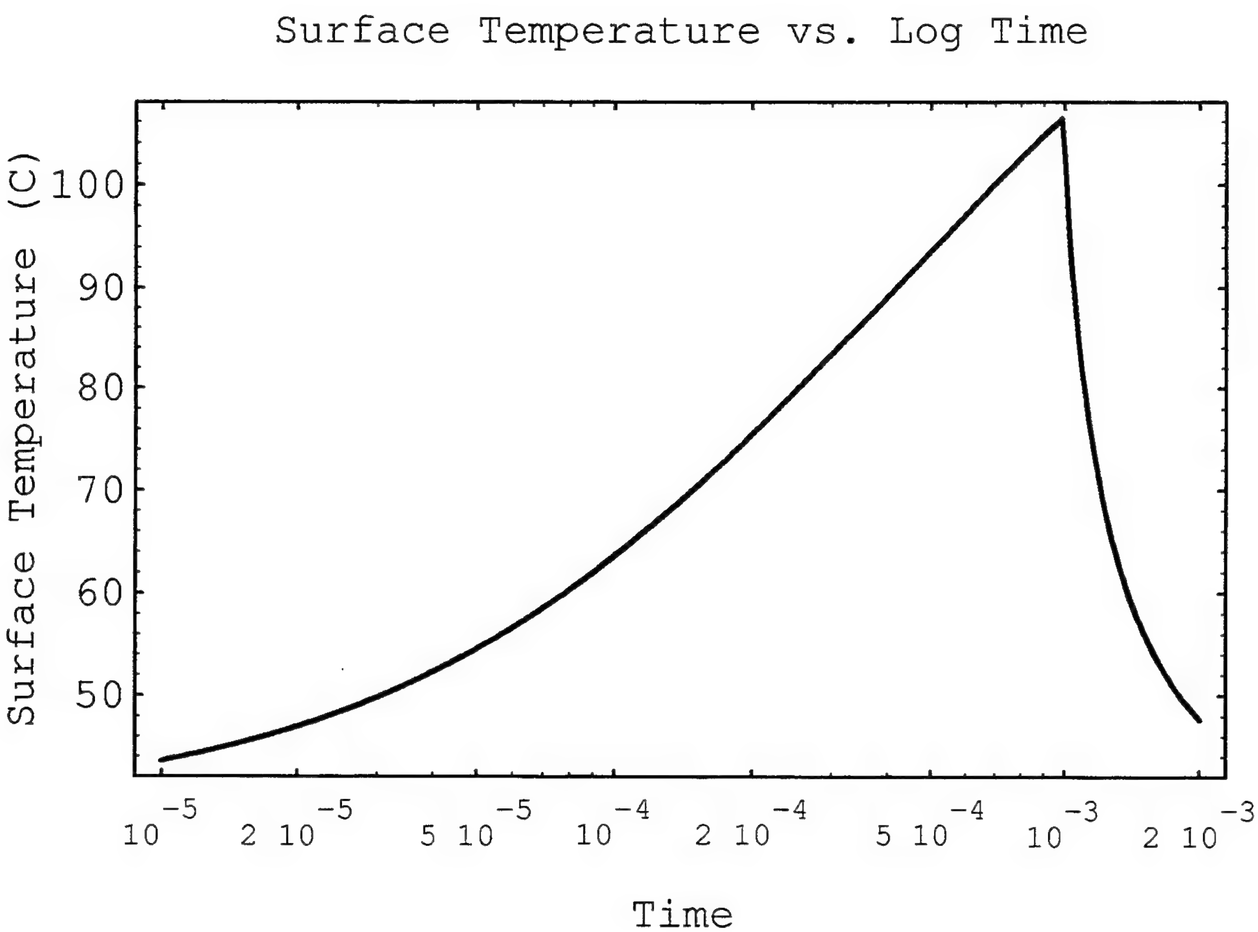

Surface Temperature vs. Log Time
Surface Temperature (C)
100
90
80
70
60
50
10^-5
2 10^-5
5 10^-5
10^-4
2 10^-4
5 10^-4
10^-3
2 10^-3
Time

Figure 5

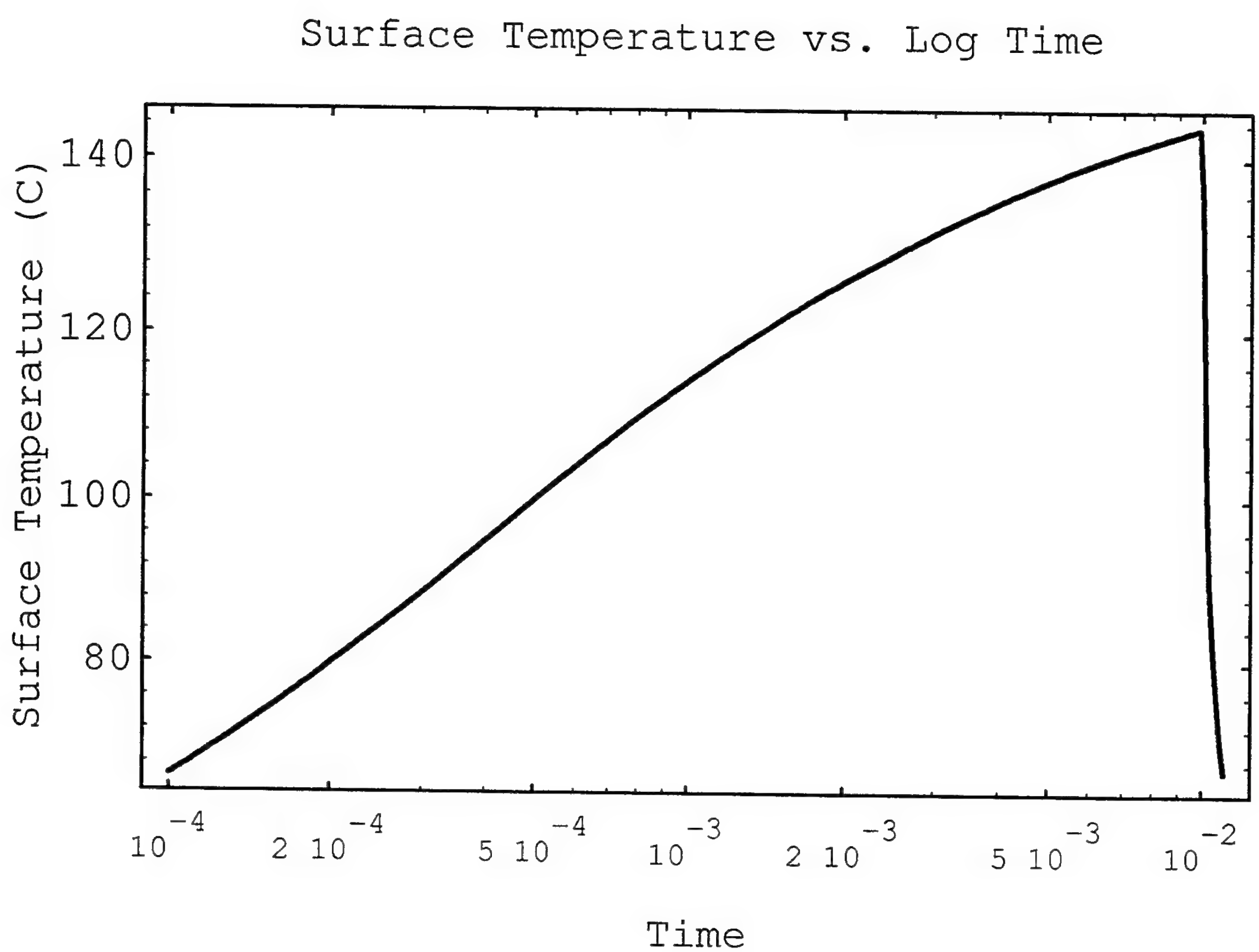

Surface Temperature vs. Log Time
Surface Temperature (C)
140
120
100
80
10^-4
2 10^-4
5 10^-4
10^-3
2 10^-3
5 10^-3
10^-2
Time

Figure 6

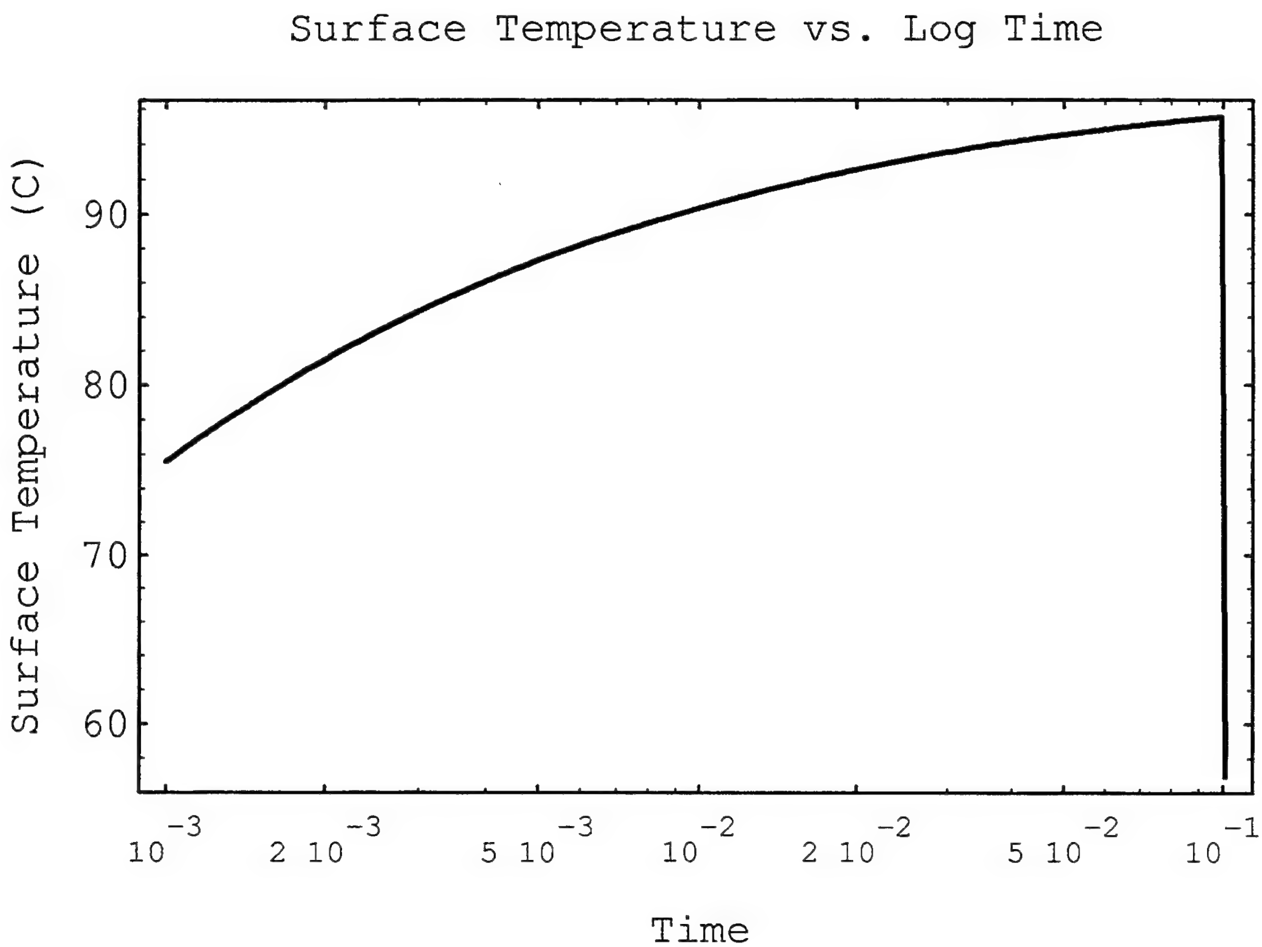
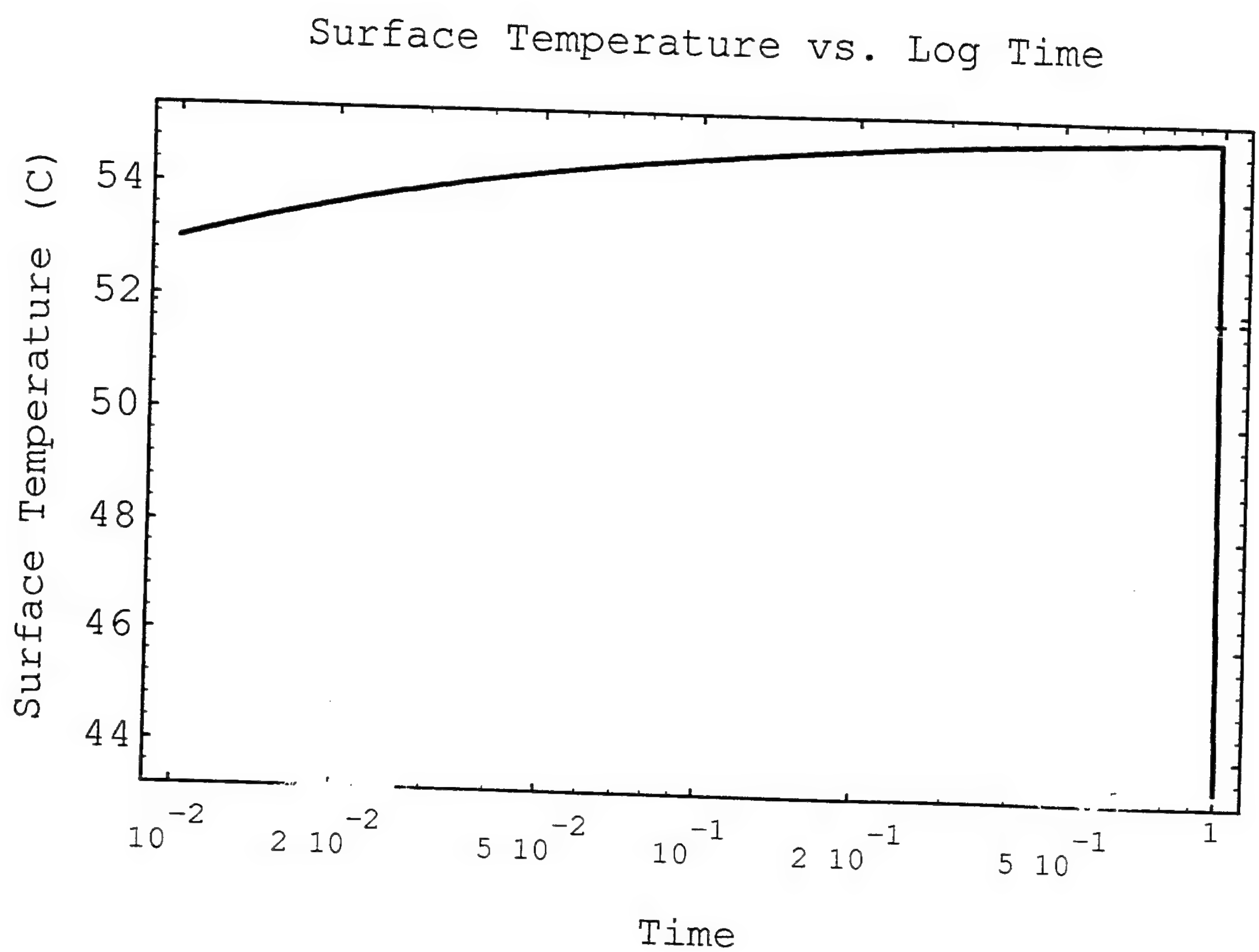

Surface Temperature vs. Log Time
Surface Temperature (C)
90
80
70
60
10^-3
2 10^-3
5 10^-3
10^-2
2 10^-2
5 10^-2
10^-1
Time

Figure 7



Surface Temperature vs. Log Time
Surface Temperature (C)
54
52
50
48
46
44
10^-2
2 10^-2
5 10^-2
10^-1
2 10^-1
5 10^-1
1
Time

Figure 8

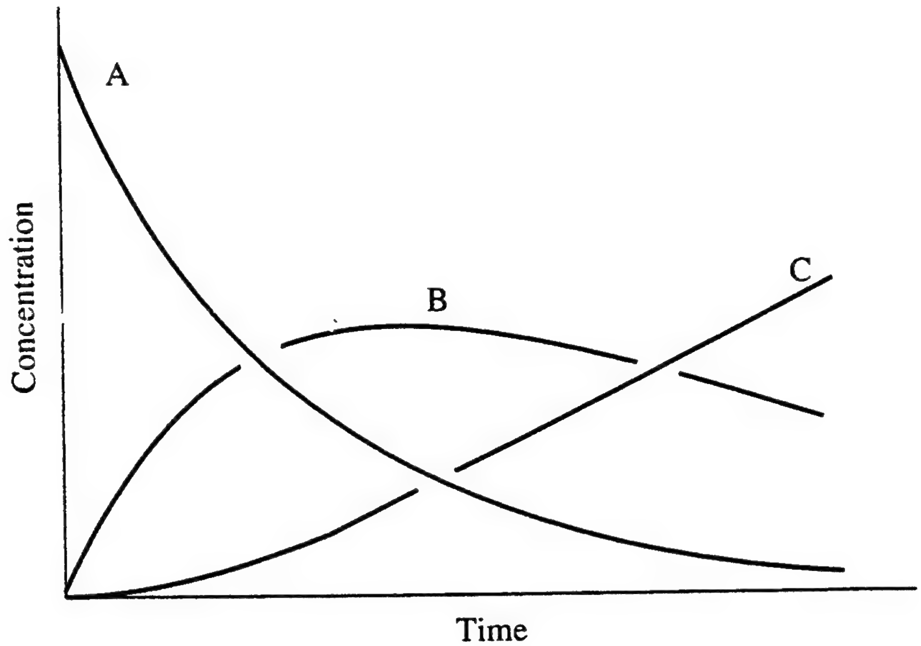


Figure 9

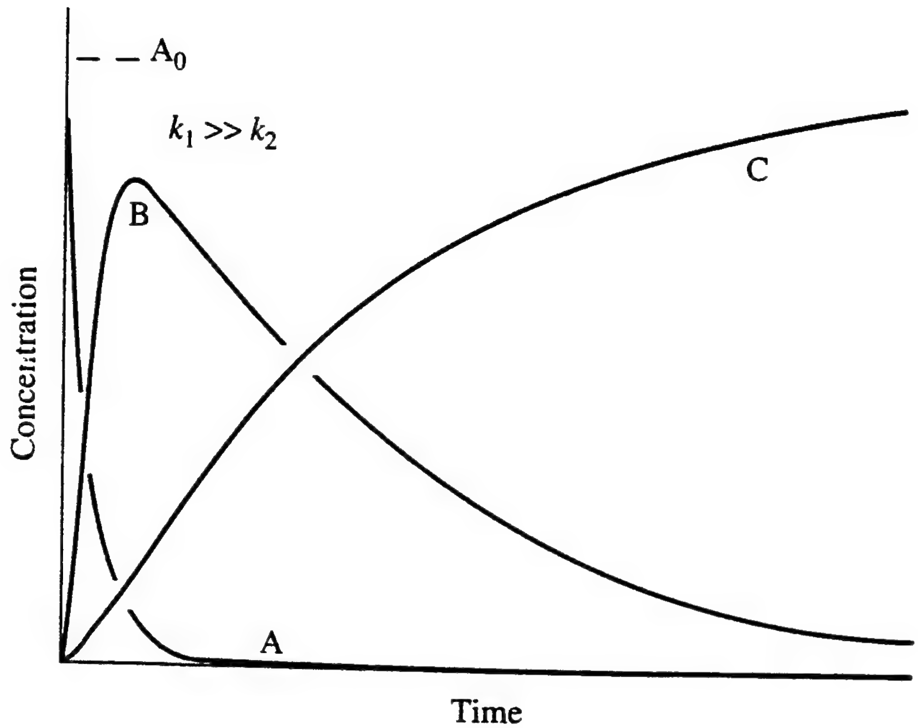


Figure 10

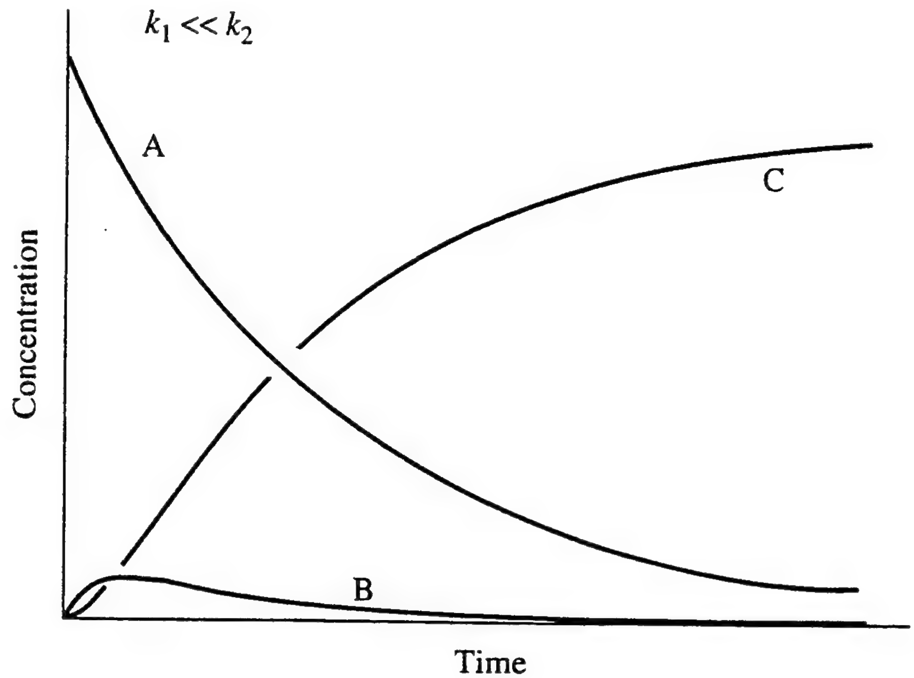


Figure 11

EFFECTS OF MENTAL WORKLOAD AND ELECTRONIC SUPPORT ON NEGOTIATION PERFORMANCE

Kenneth A. Graetz, Ph.D.
Assistant Professor
Department of Psychology
The University of Dayton
300 College Park
Dayton, OH 45469-1430

Final Report for:
Summer Research Extension Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, D.C.
and
Armstrong Laboratory

April 1996

EFFECTS OF MENTAL WORKLOAD AND ELECTRONIC SUPPORT ON NEGOTIATION PERFORMANCE

Kenneth A. Graetz, Ph.D.
Assistant Professor
Department of Psychology
The University of Dayton

Abstract

The current study used a 4-issue, variable-sum negotiation task to determine the degree to which different negotiation support system interfaces increase joint profit and reduce negotiator workload. Three displays were compared: strictly numeric, numeric with mathematical assistance, and strictly graphical. Each display type was also tested under two time constraints: 10 minutes (high time pressure) and 20 minutes (low time pressure). Although male negotiators tended to obtain higher joint profit when offered mathematical assistance, female dyads achieved their maximum joint profit with a strictly numerical display. Under time pressure, mathematically assisted dyads obtained the highest joint profit. The numerical netted the highest joint profit under low pressure conditions. Workload estimates were higher for males in the low pressure condition than for males in the high pressure condition. Overall, workload estimates did not correlate with joint profit. Potential limitations of the current workload measures and a follow-up study using a new computer interface are discussed.

EFFECTS OF MENTAL WORKLOAD AND ELECTRONIC SUPPORT ON NEGOTIATION PERFORMANCE

Kenneth A. Graetz, Ph.D.

Overview

The following report describes research and software development in the area of electronic negotiation support systems (NSS) supported by an AFOSR Summer Research Extension Grant provided to the primary author in 1995. The report discusses the following topics in order: (a) background literature regarding negotiation, cognitive processing, and mental workload, (b) the results of a laboratory experiment comparing a graphical with a text-based NSS, (c) a second laboratory study, currently in the pilot testing stage, and (d) the characteristics of Supporting Conflict Analysis and Negotiation (SCAN), a new NSS currently under development.

Negotiation and Cognition

Although considered a rational and direct approach to social conflict (Falbo, 1977; Kipnis, 1984), the process of negotiation often leads to the endorsement of suboptimal agreements and lost opportunities for more mutually beneficial outcomes. Evidence of poor performance comes from a variety of both field (Howells & Brosnan, 1972; Howells & Woodfield 1970; Lax & Sebenius, 1986; Lewicki & Litterer, 1985; Raiffa, 1982; Walton & McKersie, 1965) and laboratory studies (Pruitt & Rubin, 1986; Bazerman & Neale, 1983) and is frequently attributed to the cognitive frailty of the typical negotiator (Bazerman & Carroll, 1987; Bazerman, Magliozzi, & Neale, 1985; Bazerman & Neale, 1983; Carroll, Bazerman, & Maury, 1988; Thompson & Hastie, 1990). Specifically, researchers point toward the negotiator's use of category-based structures, rather than on-line information, to infer an opponent's interests. Although the distinction between

category-based and on-line processing in the social psychological literature resides primarily with models of person perception (Fiske & Neuberg, 1990), it applies equally well to situations involving social conflict between two or more individuals. In a negotiation, for example, on-line information might include specific details about the current dispute obtained through interaction with an opposing negotiator. On the other hand, preconceived or category-based beliefs about the opposing negotiator or the negotiation process itself may also influence the disputant's judgements and behavior.

Two such category-based structures receiving empirical attention include *fixed-pie perception* and *perceived role-obligations*. Research on the latter indicates that many individuals adopt initially tough, inflexible bargaining strategies when serving as a group's representative (e.g., Benton, 1972; Klimoski, 1972; Organ, 1971). This competitive posture is observed even in the absence of explicit group pressure and may stem from the preconception that constituencies generally desire firm, competitive representatives (Vidmar, 1971; see review by Holmes & Lamm, 1979). The second cognitive structure, fixed-pie perception, is the focus of the current studies. While not tied to a particular social role, fixed-pie perception refers to the negotiator's assumption of diametrically opposed or *zero-sum* interests (Bazerman & Neal, 1983; Thompson and Hastie, 1990). Actually, this assumption combines two, independent beliefs: (a) that the parties have opposing preferences regarding a particular issue and (b) that the value or utility associated with that issue is the same for both negotiators. Thus, the negotiator assumes that a gain by one party necessitates a loss of equal magnitude for the opposing party. While many conflicts include zero-sum issues, a large number of negotiations also include issues that afford mutually beneficial or *win-win* solutions (Walton and McKersie, 1979). Using erroneous, fixed-pie perceptions to interpret a conflict that is not strictly zero-sum may lead to unecessarily contentious bargaining

tactics (Bazerman, et. al., 1985; Carrol et. al., 1988) and a failure to unearth opportunities for mutual benefit. Evidence suggests that fixed-pie perceptions impede the discovery of preference compatibility and logrolling opportunities (Thompson & Hastie, 1990; Pruitt, 1981), thereby reducing the joint outcome obtained by the disputants. Given the potentially detrimental effects of assuming that each and every negotiation issue is strictly competitive, what factors lead negotiators to set aside their preconceptions and attend to on-line information? Conversely, what factors increase the likelihood that negotiators will utilize simplifying, category-based assumptions?

Current models of person perception claim that individuals engage in category-based processing in some situations more than in others (Fiske, 1993; Fiske & Neuberg, 1990). For example, individuals who are instructed to form accurate impressions of others (i.e., who are motivated by an *accuracy goal*) are likely to attend to individuating information (Neuberg, 1989). On the other hand, perceivers who experience increased time pressure are likely to make category-based judgements (Jamieson & Zanna, 1988; Koller & Wicklund, 1988). Past negotiation research is highly suggestive in this regard. Negotiators instructed to discuss preference information during the negotiation session (Thompson, 1991) or who expect to interact with their opponent again in the future (Heide & Miner, 1992) obtain relatively high joint outcomes. Perhaps these manipulations motivate negotiators to devote more of their cognitive resources to obtaining information about their opponents' interests. Conversely, negotiators exhibit poor performance when experiencing either increased time pressure (Carnevale & Lawler, 1987; Pruitt & Latane-Drews; 1969), increased external pressure from a constituency (Carnevale, Pruitt, & Britton, 1979; Druckman, Solomon, & Zechmeister, 1972), or increased task complexity (Kelley, Deutsch, Lanzetta, Nuttin, Shure, Faucheux, Moscovici, & Rabbie, 1970). These

variables may place additional cognitive demands on the negotiator, thereby increasing the usefulness of simplifying assumptions.

Negotiation and Mental Workload

This evidence supports a view of category-based processing in negotiation that includes some notion of cognitive demand or *mental workload*. All negotiations occur in multi-task environments in which disputants must allocate mental resources to a variety of tasks (e.g., impression management, speech comprehension, mental arithmetic, logical reasoning, and strategic planning). The concept of mental workload, already used extensively in the human factors area, is easily modified to fit the needs of negotiation researchers. In this context, mental workload refers to the portion of information processing capacity or resources that is required to meet the demands of the negotiation (cf., O'Donnell & Eggemeier, 1986; Wickens and Kramer, 1985). It is generally hypothesized that any variable that increases a negotiator's mental workload also increases the likelihood of category-based processing. Conversely, variables that reduce mental workload may either (a) reduce the need for simplifying assumptions (e.g., fixed-pie perception) or (b) allow for a redistribution of cognitive resources favoring on-line processing. Negotiators who do not need to engage in heavy amounts of mental arithmetic or who are not anxious about meeting a deadline may allocate additional cognitive resources to elaborating on information gleaned during the actual negotiation. Negotiators experiencing increased workload levels may rely heavily on category-based processing, perhaps having neither the time nor the resources to engage in a detailed analysis of on-line information. While past negotiation research suggests a negative relationship between mental workload (e.g., time pressure, task complexity) and negotiation outcomes, the hypothesized link between mental workload and category-based processing remains untested.

Recent research suggests that session-oriented NSSs, designed specifically for use by two or more negotiators interacting in a face-to-face environment, enhances the outcomes obtained by naive negotiators (see Jelassi & Foroughi, 1989; Vetschera, 1990 for a review). Most NSS include electronic tools designed to simplify, organize, and display numerical information. While several studies comparing electronically assisted and unassisted negotiations obtained favorable results (Jarke, Jelassi, & Shakun, 1987; Jelassi & Jones, 1988; Kersten, 1985), the causes of this improvement are unclear.

In some conflicts, the sheer number of negotiation issues may tax a disputant's cognitive resources. As stated by Pruitt (1981), "Human intellectual capacity severely limits the number of issues that can be considered simultaneously." (p. 14). In a study investigating the effects of mental workload on negotiation performance (Graetz, Macbeth, Seifert, & Kozar, unpublished manuscript), same-gender dyads negotiated agreements on both a 4-issue (low mental workload) and an 8-issue (high mental workload) integrative bargaining task. In addition, half of the dyads used an electronic negotiation support system (NSS) during the negotiation sessions. The results provided some evidence of the deleterious effects of mental workload on the negotiation process. As expected, increasing the number of negotiation issues significantly affected fixed-pie perception, contentious communication, and outcome satisfaction. More specifically, the hypothesis that increasing mental workload increases the likelihood of fixed-pie perception received qualified support; male estimates on the 8-issue task reflected significantly greater fixed-pie perception than their 4-issue estimates. The hypothesis that increasing mental workload increases contentious communication also obtained qualified support; electronically assisted dyads engaged in significantly more contentious communication during the 8-issue task than during the 4-issue task. Further correlational analyses revealed a significant relationship between task

complexity ratings and the proportion of contentious statements; negotiators who perceived the tasks as more complex also tended to engage in more contentious communication. Finally, the study provided support for the hypothesis that increased mental workload reduces satisfaction with negotiation outcomes. Satisfaction ratings for outcomes obtained in the 8-issue task were significantly lower than 4-issue ratings.

In a second study, Graetz et al., found that negotiation performance may depend heavily on the manner in which the task is presented. This second study presented negotiators with either a graphical representation of the negotiation task or a traditional text-based version. The discrepancy between male and female pairs with respect to overall joint total, joint total on the logrolling issues only, and fixed-pie perception was consistently larger in the text-based display condition than in graphical display condition. As expected, the performance of female negotiators appeared to improve in the graphical display condition. By de-emphasizing the numerical nature of the task, female disputants obtained a greater number of points and demonstrated less fixed-pie perception than females in the numerical display condition. Comparatively, with the exception of fixed-pie perception scores, male performance failed to improve with graphical support. While it is possible that the transition from a numerical to a graphical format reduced the apprehension of female participants, it is unclear why male performance would suffer from such a change.

A more powerful study is necessary, utilizing manipulations that illuminate possible mechanisms for the gender effects obtained in the previous studies. This formed the rationale for the study presented in this report.

Method

Participants

Seventy-three males and 73 females received $10.00 each for participating in the study. Most were undergraduate students attending summer courses at The University of Dayton. Of the 73 dyads, 11 pairs (15%) of negotiators did not reach an agreement within the time allotted. Data from these pairs were dropped from further analyses.

Negotiation Task

The 4-issue negotiation task used in the current study and presented in Table 1 was an integrative bargaining problem similar to those used and described extensively elsewhere (see Pruitt & Lewis, 1975; Thompson & Hastie, 1990; Thompson, 1991). Payoff schedules displayed five discrete levels per issue, with the negotiator's points per level listed in parentheses. The payoff schedule represented a negotiation between two a state land manager and a developer over the construction of an amusement park on state land. This 4-issue task included one compatible issue (*earliest opening day*) and one zero-sum issues (*% of gross*). Negotiators' point distributions on the compatible issue were identical. For the zero-sum issue, the two negotiators' point distributions, while equal in magnitude, reflected diametrically opposed preferences. The other two issues (*acres* and *% of in-state employees*) combined to form a logrolling pair. Each involved conflicting interests. Unlike the zero-sum issue however, each logrolling issue was less important to one negotiator than to the other. The land manager valued *% of in-state employees* over *acres*; the developer valued *acres* over *% of in-state employees*.

Independent Variables

Time pressure. Participants were allowed either 10 or 20 minutes in which to reach an agreement on all four issues. All participants were given a 2 minute warning as they approached the end of their allotted time.

Display. While negotiating agreements on the 4-issue task, participants had access to one of three computerized payoff schedules: traditional, traditional plus calculator, and graphical. The graphical computer interface was a simple database management system designed using Microsoft's Visual Basic. The computer screen included three basic displays: (a) a pull-down menu representation of the current payoff schedule, (b) a "total-points" pie chart, and (c) a vertical bar chart. While negotiating, participants could select a level for each issue by clicking the left mouse button (LMB) on the pull-down menu representing that level. As the negotiator selected levels with the LMB, green vertical bars displayed the total points per issue on the bar graph. Preexisting, black bars displayed the maximum number of points the negotiator could gain or lose for each issue. This vertical bar chart allowed negotiators to generate an offer and visually compare each offer, issue-by-issue, with the maximum point values per issue (the black bars). In addition to updating the vertical bar graph, clicking the LMB on an issue level changed the value displayed in the pie chart. This chart displayed the proportion of the total points that the negotiator would receive for the selected offer (i.e., all of the levels selected using the LMB).

Dyads in the numerical condition used a similar pull-down menu representation of the payoff schedule. However, the vertical bar chart and the pie chart were not displayed. Finally, negotiators in the numerical plus calculator condition viewed the same display as those in the numerical condition, except their total points for each offer were computed and displayed automatically.

Dyad gender. An equal number of male and female dyads participated in the study in order to assess potential gender effects. All dyads were same-gender pairs.

<u>Procedure</u>

Participants arrived in pairs. During the recruitment phase, an experimenter ensured that the dyads were unfamiliar with one another. Two experimenters, either both male or one male and one female, conducted each session. Upon arrival, participants sat at private desks to minimize pre-experimental contact and discussion. After obtaining informed consent, experimenters described the method of payment. Participants earned $10 for their participation in the study plus a chance to win a $100 bonus prize. An experimenter informed each individual that the person who earned the most points across both negotiations would receive the $100 prize and that the winner would be determined following completion of the entire study (i.e., not at the end of the experimental session). Experimenters told participants that the probability of a tie was "quite high" and that the recipient of the prize would, "...most likely be selected randomly from among the top point earners[2]".

Each negotiation involved randomly assigning participants to their respective roles. Negotiators received a role-specific *briefing sheet*, ostensibly written by their respective supervisors, that included information regarding their supervisor's preferences. The experimenter read the briefing sheet aloud, providing information about all of the issues.

The interests of the supervisor simply reflected the negotiator's own point distributions. The briefing sheets described each issue separately, excluding any comparisons across issues. Finally, each briefing sheet closed with the statement, "We understand that you may not be able to satisfy all of our requests for all four (eight) issues. We trust your judgment. Remember, it's the final package that counts. Try to negotiate a good, overall agreement." Experimenters repeated this statement immediately prior to the start of each negotiation.

Next, negotiators completed a short, paper-and-pencil exercise designed to measure their understanding of their first payoff schedule. The exercise required that negotiators compute both

the maximum and minimum total score for that negotiation as well as the total number of points that they would earn given two, randomly selected offers. Participants used the NSS to complete the exercise. The experiment required that all subjects achieve 100% accuracy on the exercise before continuing with the procedure. This prenegotiation phase (i.e., the introduction, role orientation, briefing, and exercise) took approximately 20 minutes to complete.

An experimenter then described the information displayed on the interactive screen, emphasizing the computational method and the basic meaning of each piece of information. The experimental protocol avoided encouraging negotiators to employ a particular strategy when utilizing the computer. At the end of the training session, experimenters said, "We would like you to use this tool while you negotiate, however we don't want it to interfere with any face-to-face discussion that you may have with the other negotiator."

Negotiators were allowed to communicate with one another using a computerized “chat” tool. The two computers were networked such that the tool displayed text as it was being entered by the other negotiator (and vice versa). Participants were provided with a short demonstration of the tool.

An experimenter advised negotiators as to the time limit and the consequences of not reaching an agreement. Dyads that failed to reach an agreement within the allotted time period received zero points for that negotiation. Experimenters informed the dyad that there were no limits to what could be discussed during the negotiation; the only requirement was that they never show their payoff schedules or briefing sheets to their opponents during the negotiation. Experimenters informed participants that their opponent's payoff schedule and briefing sheet included the same issues and the same levels per issue, but that some of the point distributions, "...may be different." Experimenters told dyads that a two minute warning signal would be provided. The experimenters

observed the negotiation from behind a one-way mirror, timing the duration of the negotiation with a stopwatch. Experimenters instructed participants to knock on the door of the conference room after reaching agreement on all issues.

After recording their decision, participants completed an *estimation task* and a post-negotiation questionnaire. The estimation task asked negotiators to rank order, according to their importance to the opponent, the four negotiation issues. The post-negotiation questionnaire included items designed to measure mental workload and satisfaction.

Experimenters then paid and debriefed the participants.

Dependent Variables

Total points. The total points for the dyad was computed from the final agreement. The maximum potential joint total was 15200.

Mental workload. As a measure of perceived mental workload imposed during the course of the negotiation, the NASA-TLX (Task Load Index; Hart & Staveland, 1988) was administered. This is an instrument consisting of six component subscales. The subscales measure as follows: perceived mental demand, physical demand, temporal demand, performance, effort, and frustration level. Twenty-step bipolar scales are used to obtain ratings for each subscale. A score from 1 to 20 was obtained for each subscale, and a total score, ranging from 6 to 120 was obtained for each negotiator.

The NASA-TLX originally called for a weighting procedure to be performed to determine the relative contribution of each subscale (as perceived by the participant) to the overall mental workload. Recent studies have suggested that this is not a necessary procedure and may actually severely restrict the psychometric properties of the TLX (Nygren, 1991). Through research, it has been shown that a simple, equally weighted average of the six dimensions is just as effective, if not

more so, as the old weighting procedure (Hendy, Hamilton, & Landry, 1993; Nygren, 1991). Thus, an equal weighting of the six subscales was used to obtain the overall workload score.

Satisfaction measures. Each postnegotiation questionnaire included two items related to negotiator satisfaction. Disputants rated their level of satisfaction with their obtained outcome (points) and with the overall quality of the negotiation process. The scale ranged from 1 (not at all satisfied) to 7 (very satisfied). Average scores for the dyad served as satisfaction measures.

Results

Total Points

Across all sessions, the average joint profit was 13,118.64 (SD = 1189.58). A three-way analysis of variance (ANOVA) using gender, time pressure, and display as independent variables and total points as the dependent variable yielded a significant main effect for gender, $F(1, 47) = 8.56$, $p < .01$, a significant interaction between gender and display, $F(2, 47) = 6.70$, $p < .01$, and a significant interaction between time pressure and display, $F(1, 47) = 3.30$, $p < .05$. Overall, male pairs tended to earn more points than female pairs (Ms = 13550.00 and 12822.86, respectively), however, as illustrated in Table 2, this difference was most pronounced in the calculator condition, $t(15) = 3.70$, $p < .05$, and in the graphical condition $t(14) = 2.68$, $p < .05$. There was no significant difference between males and females in the numeric condition, $t < 1.00$.

As for the pressure by display interaction, the only difference in total points between the 10 and 20 minute conditions occurred in the numerical display condition, $t(24) = -2.11$, $p < .05$. The means for total points across time pressure and display are listed in Table 3.

Mental Workload

Across all sessions, the average workload was 54.82 (SD = 10.81). A three-way analysis of variance (ANOVA) using gender, time pressure, and display as independent variables and

workload as the dependent variable yielded a significant interaction between time pressure and gender, $\underline{F}(1, 46) = 6.03$, $\underline{p} < .05$. As illustrated in Table 4, males rated the workload as significantly greater in the 20 minute versus the 10 minute condition, $t(22) = -2.38$. There was no significant workload difference across time conditions for females. Workload was not significantly correlated with total points, $r = -.002$.

Satisfaction Measures

Post-negotiation questionnaires asked participants to rate their level of satisfaction with their own outcomes and with the overall quality of the preceding negotiation. Both served as dependent variables in a four-factor MANOVA that included gender, time pressure, and display independent variables. A significant triple-order interaction was obtained, $\underline{F}(4, 90) = 2.47$, $\underline{p} < .05$. Univariate ANOVAs revealed a significant triple-order interaction for satisfaction with the overall quality of the negotiation, $\underline{F}(2, 47) = 5.34$, $\underline{p} < .01$. As illustrated in Table 5, satisfaction with the negotiation was higher in the calculator condition with a 10 versus a 20 minute time period. This disparity was greater for females versus males.

Discussion and Further Analysis

The current experiment used a dyadic, integrative bargaining task to investigate the effects of mental workload and electronic support on negotiation performance. Replicating past studies, male negotiators obtained higher joint outcomes than female negotiators, particularly when the numeric-mathematical nature of the task was emphasized (i.e., in the calculator display condition). Surprisingly, this disparity did not disappear with the transition to graphical support. Indeed, the graphical display failed to produce positive results with respect to the reduction of workload or the increase in total points. Finally, the lack of a correlation between workload and total points was unexpected and fails to replicate findings obtained in past studies.

This lack of support for the graphical task representation may be due to a number of factors. First, it is possible that the display used in the current study was too complex for the users. This does not seem likely as the workload measure would have increased had users experienced difficulties with this interface. Another possibility is that users did not use the interface, relying instead on some alternative means of offer generation (e.g., a paper and pencil). Again, this explanation can probably be ruled out as no alternative means of generating offers existed and failure to utilize the interface would have surely resulted in lower points. A more likely explanation is that the interface lacked information necessary for the effective analysis of offers. By eliminating the points in parentheses next to each level, the negotiators had no way of determining quickly the utility of counteroffers made by the opponent. In the calculator condition, users could quickly tally up the total points that would have resulted from each offer. Future versions of the graphical interface might include some graphical means of offer comparison.

Finally, findings with respect to workload are puzzling. In the current study, the TLX appeared more sensitive to the duration of the task than to the actual mental effort exerted. Thus, the insensitivity of the TLX in this particular context may necessitate the use of other workload measures in future studies. For example, physiological measures of workload may prove more sensitive to subtle changes induced by display alterations.

The SCAN Interface and a Proposal for a Second Experiment

The proposed second study will focus on the interface style of the negotiation support system (alphanumeric, mathematical help, graphical), the gender of the negotiator, and the effects of these two variables on physiological and subjective measures of mental workload (eyeblink rate, heart rate, NASA-TLX). It is expected that the use of a graphical interface style will decrease mental workload experienced by the negotiator and that this will be indicated by decreased heart

rate, increased eyeblink rate, and decreased overall workload rating on the NASA-TLX. In addition, it is expected that the graphical interface will increase the joint benefit achieved relative to those participants using the alphanumeric interface or mathematical help interface. Another of the interface styles to be studied will be termed the mathematical help interface, in which a computer based calculator is incorporated so that participants will not have to add point values in their heads. Due to the added help that this interface affords, it is expected that, relative to the alphanumeric display, it will decrease the mental workload of the negotiator (as shown by decreased heart rate, increased eyeblink rate, and decreased overall workload ratings on the NASA-TLX) and increase the joint benefit score.

It is also anticipated that there will be a difference in terms of joint benefit between the male and female negotiators. Specifically, female negotiators are expected to benefit more from the graphical display than the male negotiators. This is based on previous research that found differential effects for male and female participants (Gerhart & Rynes, 1991; Graetz et al., unpublished manuscript) in a simulated negotiation task. It is also predicted that male and female negotiators will experience differential decreases in mental workload when using the graphical interface style. Specifically, females will experience a greater decrease in workload (using the previously mentioned measures) than males when using the graphical interface style. This is based on previous research that indicated that males and females negotiate differently (Coleman, 1982; Hottes & Kahn, 1974; Gerhart & Rynes, 1991).

<u>Participants</u>

A total of 36 people will participate as partial fulfillment of an introductory psychology course requirement and each will receive $10.00 for their participation. A 2 x 3 between-subjects factorial design will manipulate negotiator gender and interface style (alphanumeric, mathematical

help, or graphical). The 18 male and 18 female participants will be randomly assigned in equal numbers to one of the three interface styles.

Negotiation Task

The task is a slight variation on one previously utilized in integrative bargaining studies (see Bazerman, Magliozzi, & Neale, 1985; Pruitt, 1981; Thompson, 1990). Participants will play the role of a theme park developer. They will be told that they are negotiating with a land manager for the state regarding the construction of a new theme park. Participants will be told to negotiate on four issues: park acreage, earliest opening date for the park, the percentage of the park's gross profit to go to the state, and the number of in-state employees to be hired by the park. Each offer will be followed by the point values (in parentheses) to the negotiator if the agreement includes that offer level (except in the graphical interface condition). This set of issues includes one compatible issue (earliest opening date), for which the point values are equal for both players. One issue is a zero-sum issue (percent gross profit), in which the point values for the two players are diametrically opposed. The remaining two issues (percent in-state employees and acres) are the issues with integrative potential. The acres issue is valued more by the park developer and percent of gross is the issue of greatest importance to the land manager. The distributive solution to the negotiation would be to agree on 90 acres of land, an opening date of 3 years, 6% of gross profit to go to the state, and 30% of the employees to be from in-state. The integrative solution would be to agree on 110 acres of land, an opening date of 1 year, 6% of gross profit to go to the state, and 50% of employees to be from in-state.

Independent Variables

Interface style. During the negotiation session, all participants will be presented with a computerized version of the payoff schedule. Three interfaces will be used. Each of the interfaces has been designed using Visual Basic 3.0 for Windows.

The negotiators in the alphanumeric condition will view the payoff schedule of offers via a series of pull-down menus. Each issue menu will pull-down to show the five possible offer levels followed by the point values associated with them in parentheses. Another menu exists to allow the negotiator to choose to make a final offer. The display will also include a manager's reply box, windows presenting counter offers from the land manager, a "send" button for the developer, and buttons corresponding to possible replies to the counter offers.

The negotiators in the mathematical help condition will view the interface used in the alphanumeric condition. In addition, two "total-points" boxes will appear on the display. As the participant clicks the left mouse button (LMB) on an offer, the total points summed over all four issues will be displayed in one of these boxes. The total points for each counter offer will also be displayed in the other box.

The third interface is a graphical representation of the offers. This NSS, called the SCAN interface (Supporting Conflict Analysis and Negotiation. Once again, the payoff schedule will be presented via a series of pull-down menus, although the point values for each offer level will not be available in parentheses. Another menu will allow the negotiator to make a final offer. The display will include the same controls that the other interfaces did, but it will also include a vertical bar chart and a pie graph.

During the process of negotiation, the participant will choose an offer level for each issue by clicking the LMB on the issue and dragging the pointer to the chosen level. As a level is chosen, a check mark will appear next to it on the menu. A blue bar will display the total points

for that level of the issue on the vertical bar chart. On the bar chart, each issue is represented by a blue bar to indicate the points awarded for the offer under consideration. Also, gray bars will appear for each issue showing the maximum number of points that could be earned for that issue. These gray bars serve as a reference point for the negotiator and will not change throughout the task. In addition, as counter offers are generated, they will be represented by red bars on the graph.

A pie graph will also be updated with every offer level that is chosen. As new offers are evaluated, the pie graph will show the relationship between the total points summed across all four issues for the offer under consideration and the maximum number of points that are possible (the whole pie). As the participant clicks the LMB on a different offer level for an issue, the bar chart and the pie graph will be updated to show the points for the offer. A second pie graph will show the proportion of points based on the manager's counter offer.

Gender. An equal number of males and females will participate. This is so potential gender effects can be studied.

Procedure

Each participant will be brought into the negotiation room. They will be asked to read and sign a form indicating their consent to participate. A general description of the task and procedure for monitoring physiological responses will be read by the researcher, along with a description of the payment method. Individuals will earn $10.00 for their participation in the study plus a chance to win a $100.00 bonus prize. The participants will be informed that the person who earns the most points totaled across all four issues would receive the prize and that the winner will be determined once the entire study is completed. Participants will be told that it is likely that more than one person will receive the same high point total and that if this is the case, the winner will be

chosen randomly from among the top point earners. Participants in the graphical interface condition will be told that they should maximize the proportion of the total pie that they receive and that awards will be made based on this amount.

The skin will be thoroughly cleansed in the areas necessary using isopropyl alcohol. Electrodes will then be placed as follows: on the superorbital notch and the infraorbital notch of the left eye, on the sphenoid bone, and on the left and right inner arm about one inch below the elbow. Once they have been properly placed, the skin resistance will be measured. If it is greater than 10,000 ohms, the skin must be recleansed and the electrodes reattached. Once this is completed, the task instructions will continue.

Participants will be then be given a briefing sheet that informs them as to what each of the issues represents and what their goal should be with regard to each issue (see Appendix C). In addition, the briefing sheet will inform the participants that they should work to try to come up with the best, overall solution that they can. Once this briefing sheet is read, instructions as to how to manipulate the interfaces will be read to the participants, allowing them the opportunity to practice using the interface and ask questions regarding its use.

Participants will be told that the computer will play the role of the land manager. They will be told that they have 10 minutes to arrive at an agreement and that if they do not agree on all four issues in that time, they will be assigned a score of zero.

Next, the participants will complete a pre-experimental exercise and questionnaire. The exercise is designed to test their understanding of the task and their priorities. They will be asked to list the offer levels for all four issues that will result in the highest and the lowest point totals for themselves. They will also be asked to evaluate two random offers to determine which is the best in terms of total points to themselves. The researcher will then check to ensure that all the

questions on this exercise are correctly answered. If they are not, then further instruction will be given. Once this is completed, they will be asked to complete a general questionnaire. This will ask questions regarding computer skills and questions regarding tobacco use and caffeine consumption. These questions will be included because the use of either of these items may affect heart rate. Once this questionnaire is completed, the negotiation will proceed.

After the negotiation is complete (or the time limit has been reached), participants will complete a post-experimental questionnaire. This will included a modified version of the NASA-TLX (Task Load Index; Hart & Staveland, 1988). The questionnaire will also ask them to evaluate their performance on a series of Likert scales. Once this is completed, the experimenter will pay and verbally debrief the participants.

The Negotiation

Each participant will be negotiating with a computer generated opponent. Therefore, some rules have been established for how the negotiation will proceed. The game proceeds as a series of offers and counter offers, with each side making preliminary offers to each other until an agreement is found which is agreeable to both and which each side determines to be the best possible solution they can reach.

The game begins when the park developer (the actual opponent) makes the first offer. To make this first offer, the participant will click the LMB on an offer level for each issue and then click the LMB on the "Send" button. At this point, the computer will evaluate the offer in relation to a predefined set of rules to determine whether it should be accepted. The first rule is that the computer will unconditionally accept the offer if it results in a point total (across all four issues) that equals or exceeds the point total for the integrative agreement (7,600). But if the point total of the offer is less than the integrative solution but greater than or equal to the point total for the

distributive solution (5,200), then the computer will give a conditional reply of "Maybe" in the reply box. The final rule applies to offers that result in point totals that are less than occurs with the distributive agreement. In this case, the reply box would return a "No" response.

Once the computer opponent has made a determination as to whether to accept or reject an offer, it then has to make a counter offer. The first offer made by the computer will always be one which maximizes the total points that the computer opponent (the land manager) could receive over all the issues (10,400). This will set a high aspiration level for the computer opponent.

After the first counter offer, the computer will use different rules to make counter offers. The simulator is programmed to emulate rational negotiators. To this end, the computer opponent will never make counter offers that result in point totals below the distributive solution point total. If the real opponent (the park developer) offers a solution that is accepted or is conditionally accepted by the computer opponent with a "Yes" or a "Maybe" in the reply box, then the computer would randomly select a counter offer that results in a point total that is greater than the point total that results from the real opponent's offer. If the real opponent makes an offer that, in terms of total points to the computer opponent, could only be beaten by one other offer, the computer opponent will make the counter offer that results in the maximum point total to the computer across all four issues. If the real opponent makes an offer that maximizes the computer's total points, then the computer opponent will simply counter with that maximal offer.

Once a counter offer is made, the park developer must respond to it. This will be done by clicking the LMB on one of the three possible response buttons. The possible responses will be: (a) Yes, I'll definitely accept this offer, (b) Yes, but I might prefer another, or (c) No.

Once a solution has been found a final offer will be made by the real opponent. If it is acceptable to the computer opponent, the negotiation will end. If not, the computer will continue to counter.

Dependent Variables

NASA-TLX (task load index). As a measure of perceived mental workload imposed during the course of the negotiation, the NASA-TLX (Task Load Index; Hart & Staveland, 1988) will be administered. This is an instrument consisting of six component subscales. The subscales measure as follows: perceived mental demand, physical demand, temporal demand, performance, effort, and frustration level. Twenty-step bipolar scales are used to obtain ratings for each subscale. A score from 0 to 100 will be obtained for each subscale, as each division of the scale represents an increment of five.

The NASA-TLX originally called for a weighting procedure to be performed to determine the relative contribution of each subscale (as perceived by the participant) to the overall mental workload. Recent studies have suggested that this is not a necessary procedure and may actually severely restrict the psychometric properties of the TLX (Nygren, 1991). Through research, it has been shown that a simple, equally weighted average of the six dimensions is just as effective, if not more so, as the old weighting procedure (Hendy, Hamilton, & Landry, 1993; Nygren, 1991). So in the proposed study, an equal weighting of the six subscales will be used to obtain the overall workload score.

Heart rate. The participant's heart rate will be recorded throughout the negotiation task using the Neuropsychological Workload Test Battery (NWTB). This is a computerized test system developed by the Armstrong Aeromedical Research Laboratory (AAMRL) for the U.S. Air Force.

An overall average heart rate will be computed. In addition, the 10 minute session will be divided into five minute increments and the average heart rate per increment will be calculated.

Eyeblink rate. The participant's eyeblink rate will be recorded throughout the negotiation using the Neuropsychological Workload Test Battery. Due to limitations of the program, the eyeblinks can only be monitored for the first 500 seconds. An overall average eyeblink rate will be computed.

Joint benefit scores. When the agreement has been reached, the actual participant (the park developer) will have earned a certain number of points over all four issues. Correspondingly, the computer (the land manager) will have also earned a certain number of points over all four issues. These two values will be added up to yield a joint benefit score. The maximal joint benefit score attainable will be 15,200 points. The higher the joint benefit score, the more the two negotiators were able to come to an integrative agreement.

Fixed-pie perception. On the post negotiation questionnaire, participants will be asked to rate of disagreement between negotiators on each of the four issues. Participants will use a 7-point Likert scale with "1" corresponding to being "very much agreed" and "7" being "very much disagreed". Although there is no direct method of measuring the amount of fixed-pie perception the participant has, the ratings of each issue relative to the other issues will allow the presence of fixed-pie perception to be detected. It is expected that those with high amounts of fixed-pie perception will rate all the issues as being on the upper end of the scale (5, 6, or 7) whereas those with little fixed-pie perception will rate the percentage of gross profit issue as a 5, 6, or 7, the opening date issue as a 1, 2, or 3, and the acreage and in-state employees issues as somewhere between 3 and 5. So in the final analysis, a comparison will be conducted among the three different conditions on the relative ratings on this question. It would be expected that the

graphical condition will show lower ratings (less fixed-pie perception) on the opening date, in-state employees.

Relevance to Air Force Goals

The current research program promises to improve negotiation support technology as well as further basic understanding of the relationship between mental effort and satisfactory negotiated agreements. Acquisition management often requires negotiation among conflicting points of view and complex requirements. Integrated process and product development stresses multidisciplinary alliances and the ability to make reasoned tradeoffs. When proved in a laboratory setting, negotiation technology can be incorporated into computer networks and meeting rooms supporting concurrent engineering. Design is both a technical and a social enterprise. Technical solutions for solving the problems of group dynamics can only benefit organizations based on teamwork and quality management.

References

Bazerman, M.H., & Carroll, J.S. (1987). Negotiator Cognition. In L.L. Cummings & B.M.Straw (Eds.). Research in Organizational Behavior. Greenwich, CT: JAI Press.

Bazerman, M.H., & Neale, M.A. (1983). Heuristics in negotiation: Limitations to dispute resolution effectiveness. In M.H. Bazerman and R.J. Lewicki (Eds) Negotiating in Organizations. Beverly Hills, CA: Sage.

Bazerman, M.H., Magliozzi, T., & Neale, M.A. (1985). The acquisition of an integrative response in a competitive market. Organizational Behavior and Human Performance, 34, 294-313.

Benton, A.A. (1972). Accountability and negotiations between group representatives. Proceedings, 80th Annual Convention of the American Psychological Association, 227-228.

Messick, D.M., & Brewer, M.B. (1983). Solving social dilemmas: A review. Review of Personality and Social Psychology, 4, 11-44.

Carnevale, P.J.D., & Lawler, D. (1986). Time pressure and the development of integrative agreements in bilateral negotiations. Journal of Conflict Resolution, 30, 636-659.

Carnevale, P.J.D., Pruitt, D.G., & Britton, S.D. (1979). Looking tough: The negotiator under constituent surveillance. Personality and Social Psychology Bulletin, 5, 118-121.

Carnevale, P.J.D., Pruitt, D.G., & Seilheimer, S. (1981). Looking and competing: Accountability and visual access in integrative bargaining. Journal of Personality and Social Psychology, 40, 111-120.

Carroll, J.S., Bazerman, M.H., & Maury, R. (1988). Negotiator cognitions: A descriptive approach to negotiators' understanding of their opponents. Organizational behavior and Human Decision Processes, 41, 352-370.

Coleman, A. (1982). Game theory and experimental games. New York: Pergamon Press.

Deutsch, M. (1993). Educating for a peaceful world. American Psychologist, 48, 510-517.

Druckman, D., Solomon, D., & Zechmeister, K. (1972). Effect of representational role obligations on the process of children's distribution of resources. Sociometry, 35, 387-410.

Falbo, T. (1977). The multidemensional scaling of power strategies. Journal of Personality and Social Psychology, 35, 537-548.

Fiske, S.T., & Neuberg, S.L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. Advances in Experimental Social Psychology, 23, 1-74.

Gerhart, B., & Rynes, S. (1991). Determinants and consequences of salary negotiations by male and female MBA graduates. Journal of Applied Psychology, 76, 256-262.

Heide, J.B., & Miner, A.S. (1992). The shadow of the future: Effects of anticipated interaction and the frequency of contact on buyer-seller cooperation. Academy of Management Journal, 35, 265-291.

Holmes, J.G., & Lamm, H. (1979). Boundary roles and the reduction of conflict. In W.G. Austin and S. Worchel (Eds.), The social psychology of intergroup relations. (pp. 304-317). Monterey CA:Brooks/Cole.

Hottes, J. & Kahn, A. (1974). Sex differences in a mixed-motive conflict situation. Journal of Personality, 42, 260-275.

Jamieson, D.W., & Zanna, M.P. (1988). Need for structure in attitude formation and expression. In A.R. Pratkanis, S.J. Breckler, and A.G. Greewald (Eds) Attitude Structure and Function, pp. 383-406. Hillsdale, NJ: Erlbaum.

Jarke, M., Jelassi, M.T., & Shakun, M.F. (1987). Mediator: Towards a negotiation support system. European Journal of Operational Research, 31, 314-334.

Jelassi, M.T., & Foroughi, A. (1989). Negotiation support systems: An overview of design issues and existing software. Decision Support Systems: The International Journal, Special Issue on Group Decision Support Systems, June.

Jelassi, M.T., & Jones, B.H. (1988). Getting to yes with NSS: How computers can support negotiations. In R.M. Lee, A.M. McCosh, and P. Migliarese (Eds.), Organizational decision support systems, Amsterdam, North Holland, pp. 75-85.

Kelley, H.H., Deutsch, M., Lanzetta, J.T., Nuttin, J.M., Shure, G.H., Faucheux, C., Moscovici, S., & Rabbie, J.M. (1970). A comparative experimental study of negotiation behavior. Journal of Personality and Social Psychology, 16, 411-438.

Kelly, J.R., & Karau, S.J. (1993). Entrainment of creativity in small groups. Small Group Research, 24, 179-198.

Kersten, G. (1985). NEGO: Group decision support system. Information and Management, 8, 237-246.

Kipnis, D. (1984). The use of power in organizations and interpersonal settings. In S. Oskamp (Ed.), Applied Social Psychology Annual (Vol. 5, pp. 179-210). Newbury Park, CA: Sage.

Klimoski, R.J. (1972). The effects of intragroup forces on intergroup conflict resolution. Organizational Behavior and Human Performance, 8, 363-383.

Kolb, D.M., & Coolidge, G.C. (1991). Her place at the table: A consideration of gender issues in negotiation. In J.W. Breslin & J.Z. Rubin (Eds.), Negotiation theory and practice (pp. 261-277). Boston: Program on Negotiation Books at Harvard Law School.

Lax, D.A., & Sebenius, J.K. (1986). The manager as negotiator. New York: Free Press.

Lewicki, R., & Litterer, J. (1985). Negotiation. Homewood, IL: Richard D. Irwin, Inc.

Neuberg, S.L. (1989). The goal of forming accurate impressions during social interactions: Attenuating the impact of negative expectancies. Journal of Personality and Social Psychology, 56, 374-386.

O'Donnell, R.D., & Eggemeier, F.T. (1986). Workload assesment methodology. In K.R. Boff, L. Kaufman, & J. Thomas (Eds.) Handbook of Perception and Human Performance, Vol. 2: Cognitive Processes and Performance, pp. 42/1-42/49, New York: John Wiley and Sons.

Organ, D.W. (1971). Some variables affecting bourdary role behavior. Sociometry, 34, 524-537.

Pruitt D.G., & Drews, J.L. (1969). The effect of time pressure, time elapsed, and the opponent's concession rate on behavior in negotiation. Journal of Experimental Social Psychology, 5, 43-60.

Pruitt, D.G., & Lewis, S.A. (1975). Development of integrative solutions in bilateral negotiation. Journal of Personality and Social Psychology, 31, 621-633.

Pruitt, D.G., & Rubin, J.Z. (1986). Social conflict: Escalation, stalemate, and settlement. New York: Random House.

Pruitt, D.G. (1981). Negotiation behavior. New York: Academic Press.

Raiffa, H. (1982). The art and science of negotiation. Cambridge, MA: Harvard University Press.

Reid, G.B., & Nygren, T.E. (1988). The subjective workload assessment technique: A scaling procedure for measuring mental workload. In P.A. Hancock and N. Meshkati (Eds.) Human Mental Workload, pp. 185-218, Amsterdam: North Holland.

Rifkin, J. (1984). Mediation form a feminist perspective: Problems and promise. Law and Inequality, 21, 2.

Thompson, L.L., & Hastie, R. (1990). Social Perception in negotiation. Organizational Behavior and Human Decision Processes, 47, 98-123.

Thompson, L.L. (1991). Information exchange in negotiation. Journal of Experimental Social Psychology, 27, 161-179.

Thompson, L.L. (1990). The influence of experience on negotiation performance. Journal of Experimental Social Psychology, 26, 528-544.

van de Kragt, A.J.C., Orbell, J.M., & Dawes, R.M. (1983). The minimal contributing set as a solution to public goods problems. American Political Science Review, 77, 112-122.

Vidmar, N. (1971). Effects of representational roles and mediators on negotiation effectiveness. Journal of Personality and Social Psychology, 17, 48-58.

Walton, R.E., & McKersie, R.B. (1965) A behavioral theory of labor negotiations: An analysis of a social interaction system. New York: McGraw-Hill.

Wickens, C.D., & Kramer, A.F. (1985). Engineering psychology. Annual Review of Psychology, New York: Annual Reviews, Inc.

Table 1

Developer and Land Manager Payoff Schedules for the Negotiation Task

Developer's Payoff Schedule			
Acres	Earliest Opening Day	% of Gross	% of In-State Employees
110 Acres (4000)	1 Year (2400)	2% (2400)	10% (1600)
100 Acres (3000)	2 Years (1800)	4% (1800)	20% (1200)
90 Acres (2000)	3 Years (1200)	6% (1200)	30% (800)
80 Acres (1000)	4 Years (600)	8% (600)	40% (400)
70 Acres (0)	5 Years (0)	10% (0)	50% (0)

Manager's Payoff Schedule			
Acres	Earliest Opening Day	% of Gross	% of In-State Employees
110 Acres (0)	1 Year (2400)	2% (0)	10% (0)
100 Acres (400)	2 Years (1800)	4% (600)	20% (1000)
90 Acres (800)	3 Years (1200)	6% (1200)	30% (2000)
80 Acres (1200)	4 Years (600)	8% (1800)	40% (3000)
70 Acres (1600)	5 Years (0)	10% (2400)	50% (4000)

Note. Points for each issue are listed in parentheses. All payoff schedules used in the study presented levels for both players arranged from most to least preferred.

Table 2

Average Joint Profit Across Dyad Gender and Display Type

	Display Type		
Total Points	Numeric	Calculator	Graphical
Male Dyads	12,780	14,325	13,800
Female Dyads	13,162	12,688	12,400

Table 3

Average Joint Profit Across Time Pressure and Display Type

	Display Type		
Total Points	Numeric	Calculator	Graphical
10 Minutes	12,579	13,767	13,200
20 Minutes	13,454	13,291	12,650

Table 4

Average Workload Across Dyad Gender and Time Pressure

	Dyad Gender	
Workload	Male	Female
10 Minutes	47.75	56.41
20 Minutes	58.46	54.50

Table 5

Average Satisfaction With Negotiation Across Dyad Gender, Display Type, and Time Pressure

	Display Type		
Satisfaction	Numeric	Calculator	Graphical
Male Dyads			
10 Minutes	5.35	4.83	5.67
20 Minutes	5.50	5.80	5.83
Female Dyads			
10 Minutes	4.72	6.50	4.70
20 Minutes	5.50	4.83	6.40

REGRESSION TO THE MEAN IN HALF-LIFE STUDIES

Pushpa L. Gupta
Professor
Department of Mathematics & Statistics

University of Maine
Orono, ME 04469

Final Report for:
Summer Research Extension Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

and

University of Maine

December 1995

REGRESSION TO THE MEAN IN HALF-LIFE STUDIES

Pushpa L. Gupta
Professor
Department of Mathematics & Statistics
University of Maine

ABSTRACT

Half-life studies of biomarkers for environmental toxins in humans are generally restricted to a few measurements per subject taken at least one half-life after exposure. The initial dose is usually unknown because the exposure occurred before the substance was known to be toxic. In this setting, subjects are selected for inclusion in the study if their measured body burden is above a threshold (C), determined by the distribution of the biomarker in a control population. We assume a simple one-compartment first order decay model and a log-normal biomarker distribution, which together imply a repeated measures linear model relating the logarithm of the biomarker and time, with the slope being the negative of the decay rate (λ). Unless the data set is properly conditioned, we show that ordinary weighted least squares estimates of λ are biased due to regression toward the mean. Formulae are presented in the special case that 3 measurements per subject are available. Generalizations to k measurements per subject are straightforward.

These results are applied to a half-life study of 2, 3, 7, 8 tetrachlorodibenzo-p-dioxin (dioxin) in veterans of operation Ranch Hand.

REGRESSION TO THE MEAN IN HALF LIFE STUDIES

Pushpa L. Gupta

1. INTRODUCTION

Half-life studies of biomarkers for environmental toxins in humans are generally restricted to a few measurements per subject taken at least one half-life after exposure. The initial dose is usually unknown because the exposure occurred before the substance was known to be toxic. In this setting, subjects are selected for inclusion in the study if their measured body burden is above a threshold (C), determined by the distribution of the biomarker in a control population. We assume a simple one-compartment first order decay model and a log-normal biomarker distribution, which together imply a repeated measures linear model relating the logarithm of the biomarker and time, with the slope being the negative of the decay rate (λ). Unless the data set is properly conditioned, we show in section 2 that ordinary weighted least squares estimates (WLSE) of λ are biased due to regression toward the mean (James (1973); Senn & Brown (1985)). If the within-subject correlation matrix is banded, we show that the weighted least squares estimate of λ is unbiased when the data set is conditioned on all repeated measures being above a line with slope $-\lambda$. In particular if the within-subject correlation matrix is auto regressive of order 1 or has compound symmetry, then the unbiasedness of WLSE of λ is automatically

satisfied. These results are applied to a half-life study of 2, 3, 7, 8 tetrachlorodibenzo-p-dioxin (dioxin) in veterans of operation Ranch Hand. In section 3 we use the dioxin data to study the bias, variance and mean squared error (MSE), of the estimates. Results are displayed in Tables 1-4. Finally, in section 4 some conclusions and recommendations are presented for future work.

2. BIAS IN THE ESTIMATION OF DECAY RATE

We have assumed that a single exposure produced an elevation of the TCDD body burden above background level and that the first-order kinetics model

$$C_t=C_o e^{-\lambda t}. \tag{2.1}$$

holds, where C_t is the TCDD concentration t years after exposure measured in parts per trillion, C_o is the (unknown) initial exposure, and λ is a constant but unknown decay rate. Based on (2.1), the true population half-life is $t_{1/2} = \frac{\ln 2}{\lambda}$. If we take the natural logarithm of (2.1) we obtain

$$\ln C_t = \ln C_o - \lambda t. \tag{2.2}$$

Thus, (2.2) can be regarded as a motivating equation for a model which can accommodate multiple measurements per subject as well as covariates. Such a model is known as a fixed subject-effects model with repeated measures and is described below:

$$Y_{ij} = \beta_0 + \tau_i + \beta_1 t_{ij} + \epsilon_{ij},\ j = 1, 2, 3;\ i = 1, 2, \ldots, n, \tag{2.3}$$

where Y_{ij} represents the natural logarithm of the jth background

corrected TCDD measurement on the ith subject t_{ij} years after exposure, $-\beta_1$ represents the common decay rate λ, β_0 represents the average intercept for all subjects, τ_i represents the fixed subject effect for the ith subject, and ϵ_{ij} is the residual error term for Y_{ij}. Let $Y_i = \begin{bmatrix} Y_{i1} \\ Y_{i2} \\ Y_{i3} \end{bmatrix}$ be the observation vector for the ith individual such that

$$Y_i \sim N_3(X_i\beta, \Sigma), \; i = 1, 2, \ldots, n.$$

Here X_i is a 3x(n+2) matrix of known design indicators given by

$$X_i = \left[\begin{matrix} 1 & t_{i1} & 1 \\ 1 & t_{i2} & 1 \\ 1 & t_{i3} & 1 \end{matrix} \middle| \; 0_{3x(n-1)} \right] E_i,$$

E_i is an elementary matrix obtained from the (n+2) x (n+2) identity matrix by the interchange of 3rd and (i+2)th columns, $\beta = (\beta_0 \; \beta_1 \; \tau_1 \; \tau_2 \; \ldots \; \tau_n)'$ is a (n+2) x 1 vector of unknown regression coefficients, and $\Sigma = \begin{bmatrix} \theta_0 & \theta_1 & \theta_2 \\ \theta_1 & \theta_0 & \theta_3 \\ \theta_2 & \theta_3 & \theta_0 \end{bmatrix}$ is the unknown variance-covariance matrix with equal variances. We can write the combined model for all of the data in a matrix form by letting

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \qquad X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}, \qquad \text{and } V = diag(\Sigma, \Sigma, \ldots, \Sigma)_{(3n)x(3n)}.$$

Then the model for the entire observation vector is

$$Y \sim N\,(X\beta,\ V). \tag{2.4}$$

When Σ is known, the generalized least-squares estimator β is found by minimizing the quadratic form $\sum_{i=1}^{n} Q_i(\beta,\Sigma)$, where

$$Q_i(\beta,\Sigma) = (Y_i - X_i\beta)^T \Sigma^{-1} (Y_i - X_i\beta).$$

The solution is $\hat{\beta} = (X'\ V^{-1}\ X)^{-1}\ X'\ V^{-1}\ Y$.

When Σ is unknown and therefore V is unknown, which is usually the case in practice, V is replaced by its estimator S, so then $\hat{\beta} = (X'\ S^{-1}\ X)^{-1}\ X'\ S^{-1}\ Y,$. Under the normal distribution assumption $\hat{\beta}$ is also the maximum likelihood estimator of β.

In this section we would first like to find the estimator of β and in turn the estimator $\hat{\beta}_1$ of β_1 and then find the expected value of $\hat{\beta}_1$ under truncation, i.e. when $(Y_{i1}\ Y_{i2}\ Y_{13}) > (C_1\ C_2\ C_3)$, where C_1, C_2, C_3 are known constants. The weighted least squares estimate of the vector of parameters β (for model (2.4)) is given by

$$\hat{\beta} = (X'\ V^{-1}X)^{-1}\ X'V^{-1}Y,$$

$$X' \, V^{-1} \, X = \frac{1}{|\Sigma|} \left[\begin{array}{cc|cccc} na_{11} & \sum_{i=1}^{n} a_{21}^{(i)} & a_{11} & a_{11} & \ldots & a_{11} \\ \sum_{i=1}^{n} a_{21}^{(i)} & a_{22} & a_{21}^{(1)} & a_{21}^{(2)} & \ldots & a_{21}^{(n)} \\ \hline a_{11} & a_{21}^{(1)} & & & & \\ a_{11} & a_{21}^{(2)} & & a_{11} I_n & & \\ \vdots & \vdots & & & & \\ a_{11} & a_{21}^{(n)} & & & & \end{array} \right], \tag{2.5}$$

$$X'V^{-1}Y = \frac{1}{|\Sigma|} \begin{bmatrix} \sum_{j=1}^{3} b_{1j} S_{yj} \\ \sum_{i=1}^{n} \sum_{j=1}^{3} \; b_{2j}^{(i)} \; Y_{ij} \\ b_{11} Y_{11} + b_{12} Y_{12} + b_{13} Y_{13} \\ \vdots \\ b_{11} Y_{i1} + b_{12} Y_{i2} + b_{13} Y_{i3} \\ \vdots \\ b_{11} \; Y_{n1} + b_{12} \; Y_{n2} + b_{13} Y_{n3} \end{bmatrix}. \tag{2.6}$$

The following notations have been used above:

$a_{21}^{(i)} = t_{i1} \, b_{11} + t_{i2} \, b_{12} + t_{i3} \, b_{13}$, where

$b_{11} = (\theta_0 - \theta_3) \; (\theta_0 + \theta_3 - \theta_1 - \theta_2)$,

$b_{12} = (\theta_0 - \theta_2) \; (\theta_0 + \theta_2 - \theta_1 - \theta_3)$,

$b_{13} = (\theta_0 - \theta_1) \; (\theta_0 + \theta_1 - \theta_2 - \theta_3)$,

$a_{11} = b_{11} + b_{12} + b_{13}$,

$b_{21}^{(i)} = t_{i1} \; (\theta_0^2 - \theta_3^2) \; + t_{i2} \; (\theta_2 \, \theta_3 \; - \theta_0 \, \theta_1) \; + t_{i3}(\theta_1 \, \theta_3 - \theta_0 \, \theta_2)$,

$\qquad = b_{11} t_i + \Delta \, (\theta_1 \theta_3 - \theta_0 \theta_2 - \theta_0^2 + \theta_3^2)$, for equal Δ,

$$b_{22}^{(i)} = t_{i1}\ (\theta_2\ \theta_3 - \theta_0\ \theta_1) + t_{i2}\ (\theta_0^2 - \theta_2^2) + t_{i3}(\theta_1\ \theta_2 - \theta_0\ \theta_3)$$

$$= b_{12}t_i + \Delta(\theta_1\theta_2 - \theta_0\theta_3 - \theta_2\theta_3 + \theta_0\theta_1)\ ,\ \text{for equal } \Delta\ ,$$

$$b_{23}^{(i)} = t_{i1}\ (\theta_1\ \theta_3 - \theta_0\ \theta_2) + t_{i2}\ (\theta_1\ \theta_2 - \theta_0\ \theta_3) + t_{i3}\ (\theta_0^2 - \theta_1^2)$$

$$= b_{13}t_i + \Delta(\theta_0^2 - \theta_1^2 - \theta_1\theta_3 + \theta_0\theta_2)\ ,\ \text{for equal } \Delta\ ,$$

$$a_{21}^{(i)} = t_{i1}b_{11} + t_{i2}b_{12} + t_{i3}b_{13} = b_{21}^{(i)} + b_{22}^{(1)} + b_{23}^{(1)}$$

$$= a_{11}\ t_i + \Delta(b_{13} - b_{11})\ ,\quad \text{for equal } \Delta\ ,$$

$$a_{22} = \sum_{i=1}^{n}\ t_{i1}\ b_{21}^{(i)} + t_{i2}\ b_{22}^{(i)} + t_{i3}\ b_{23}^{(i)}\ ,$$

$$|\Sigma| = \theta_0\ (\theta_0^2 - \theta_1^2 - \theta_2^2 - \theta_3^2) + 2\theta_1\ \theta_2\ \theta_3\ ,$$

$$\Sigma^{-1} = \frac{1}{|\Sigma|}\begin{bmatrix} \theta_0^2 - \theta_3^2 & \theta_2\ \theta_3 - \theta_0\ \theta_1 & \theta_1\ \theta_3 - \theta_0\ \theta_2 \\ \theta_2\ \theta_3 - \theta_0\ \theta_1 & \theta_0^2 - \theta_2^2 & \theta_1\ \theta_2 - \theta_0\ \theta_3 \\ \theta_1\ \theta_3 - \theta_0\ \theta_2 & \theta_1\ \theta_2 - \theta_0\ \theta_3 & \theta_0^2 - \theta_1^2 \end{bmatrix},$$

and

$$s_{yj} = \sum_{i=1}^{n}\ Y_{ij},\quad j = 1,2,3. \tag{2.7}$$

In order to find $\hat{\beta}_1$ = [2*nd* row of $(X'V^{-1}X)^{-1}]X'V^{-1}Y$, first we need to find the 2nd row of $(X'V^{-1}X)^{-1}$. Second row of $(X'V^{-1}X)^{-1}$ can be obtained by considering a generalized inverse of $X'V^{-1}X$ since it is singular. For obtaining the generalized inverse we first drop the last column and last row of $X'V^{-1}X$ and write the

remaining matrix as $\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$, where

$$A_{11} = \frac{1}{|\Sigma|}\begin{bmatrix} n\, a_{11} & \sum_{i=1}^{n} a_{21}^{(i)} \\ \sum_{i=1}^{n} a_{21}^{(i)} & a_{22} \end{bmatrix}, \qquad A_{12} = \frac{1}{|\Sigma|}\begin{bmatrix} a_{11} & a_{11} & \dots & a_{11} \\ a_{21}^{(1)} & a_{21}^{(2)} & \dots & a_{21}^{(n-1)} \end{bmatrix}$$

$A_{22} = \frac{a_{11}}{|\Sigma|} I_{n-1}$, *and* $A_{21} = A_{12}'$.

Consider the following:

(i) $$(A_{11} - A_{12}\, A_{22}^{-1}\, A_{21})^{-1} = \frac{|\Sigma|}{D}\begin{bmatrix} a_{22} - \frac{1}{a_{11}}\sum_{i=1}^{n} (a_{21}^{(i)})^2 & -a_{21}^{(n)} \\ -a_{21}^{(n)} & a_{11} \end{bmatrix}$$

(ii) $$-(A_{11} - A_{12}\, A_{22}^{-1}\, A_{21})^{-1} A_{12}\, A_{22}^{-1} =$$

$$\frac{-|\Sigma|}{a_{11} D}\begin{bmatrix} \longleftarrow\text{---------------}\textit{first row}\text{---------------}\longrightarrow \\ a_{11}(a_{21}^{(1)} - a_{21}^{(n)}) \;\; a_{11}(a_{21}^{(2)} - a_{21}^{(n)}) \dots a_{11}(a_{21}^{(n-1)} - a_{21}^{(n)}) \end{bmatrix}$$

where $D = a_{11}\, a_{22} - \sum_{i=1}^{n} (a_{21}^{(i)})^2$.

The second row of $(X'\, V^{-1}\, X)^{-1}$ amounts to finding the second row of the matrix

$$[(A_{11} - A_{12}\; A_{22}^{-1}\, A_{21})^{-1} \quad -(A_{11} - A_{12}\, A_{22}^{-1}\, A_{21})^{-1}\, A_{12}\, A_{22}^{-1} \quad 0]$$

$$= \frac{|\Sigma|}{D}\begin{bmatrix} \longleftarrow\text{--------------------}\textit{first row}\text{--------------------}\longrightarrow \\ -a_{21}^{(n)} \;\; a_{11} \;\; -(a_{21}^{(1)} - a_{21}^{(n)}) \;\; -(a_{21}^{(2)} - a_{21}^{(n)}) \dots \;\; -(a_{21}^{(n-1)} - a_{21}^{(n)})\; 0 \end{bmatrix}$$

Thus $\hat{\beta}_1$ = [2*nd* row of $(X'\, V^{-1}\, X)^{-1}$] $X' V^{-1}\, Y$

$$= \frac{1}{D}\sum_{i=1}^{n}\left\{\begin{array}{l}[a_{11}\,b_{21}^{(i)} - b_{11}\,a_{21}^{(i)}]\,Y_{i1} + [a_{11}\,b_{22}^{(i)} - b_{12}\,a_{21}^{(i)}]\,Y_{i2} \\ + [a_{11}\,b_{23}^{(i)} - b_{13}\,a_{21}^{(i)}]\,Y_{i3}\end{array}\right\}. \tag{2.8}$$

As a special case, let

$$\begin{pmatrix} Y_{i1} \\ Y_{i2} \\ Y_{i3} \end{pmatrix} \sim N_3\left(\begin{pmatrix} \beta_0+\beta_i t_{i1}+\tau_i \\ \beta_0+\beta_1 t_{i2}+\tau_i \\ \beta_0+\beta_1 t_{i3}+\tau_i \end{pmatrix}, \Sigma\right),\ i = 1, 2, \ldots, n,$$

where Σ is an equivariance matrix specified earlier.

We have introduced this model to study the effect of regression to the mean (James (1973); Senn & Brown (1985)). It may be mentioned that $\lambda = \tau/\Delta (= -\beta_1$ in the repeated measures model (2.4)). Then

$E(\hat{\beta}_1 | Y_1 > C_1,\ Y_2 > C_2,\ Y_3 > C_3)$

$$= \frac{\sum_{i=1}^{n}}{D}\left[\begin{array}{l}(a_{11}\,b_{21}^{(i)} - b_{11}\,a_{21}^{(i)})\,E(Y_{i1}|Y_{i1}>C_1, Y_{i2}>C_2, Y_{i3}>C_3) \\ + (a_{11}b_{22}^{(i)} - b_{12}a_{21}^{(i)})\,E(Y_{i2}|Y_{i1}>C_1, Y_{i2}>C_2, Y_{i3}>C_3) \\ (a_{11}\,b_{23}^{(i)} - b_{13}\,a_{21}^{(i)})\,E(Y_{i3}|Y_{i1}>C_1, Y_{i2}>C_2, Y_{i3}>C_3)\end{array}\right.$$

$$= \frac{\sum_{i=1}^{n}}{D}\left[\begin{array}{l}(a_{11}b_{21}^{(i)} - b_{11}a_{21}^{(i)})\,[\beta_0 + \beta_1 t_{i1} + \tau_i + \frac{\sqrt{\theta_0}}{\alpha_i}E(X_{i1}|X_{i1}>a_{i1}, X_{i2}>a_{i2}, X_{i3}>a_{i3}) \\ + (a_{11}b_{22}^{(i)} - b_{12}a_{21}^{(i)})\,[\beta_0 + \beta_1 t_{i2} + \tau_i + \frac{\sqrt{\theta_0}}{\alpha_i}E(X_{i2}|X_{i1}>a_{i1}, X_{i2}>a_{i2}, X_{i3}>a_{i3}) \\ + (a_{11}b_{23}^{(i)} - b_{13}a_{21}^{(i)})\,[\beta_0 + \beta_1 t_{i3} + \tau_i + \frac{\sqrt{\theta_0}}{\alpha_i}E(X_{i3}|X_{i1}>a_{i1}, X_{i2}>a_{i2}, X_{i3}>a_{i3})\end{array}\right]$$

For equal Δ,

$E(\hat{\beta}_1 | Y_1 > C_1,\ Y_2 > C_2,\ Y_3 > C_3)$

$$= \frac{\sum_{i=1}^{n}}{D}\left[\begin{array}{l}(\beta_0+\beta_1 t_i+\tau_i)\,\{(a_{11}b_{21}^{(i)} - b_{11}a_{21}^{(i)}) + (a_{11}b_{22}^{(i)} - b_{12}a_{21}^{(i)}) + (a_{11}b_{23}^{(i)} - b_{13}a_{21}^{(i)})\} \\ + \Delta\beta_1\{(a_{11}\,(b_{23}^{(i)} - b_{21}^{(i)}) - a_{21}^{(i)}\,(b_{13} - b_{11})\} + \frac{\sqrt{\theta_0}}{\alpha_i}\,(*)\end{array}\right]$$

$$= \frac{\sum_{i=1}^{n}}{D}\left[\begin{array}{l}(\beta_0+\beta_1 t_i+\tau_i)(0)+\Delta^2\beta_1\{a_{11}(2\theta_0^2-\theta_1^2-\theta_3^2-2\theta_1\theta_3+2\theta_0\theta_2)-(b_{11}-b_{12})^2\} \\ +\frac{\sqrt{\theta_0}}{\alpha_i}(*)\end{array}\right].$$

$$(2.9)$$

$= \beta_1 + \frac{\sqrt{\theta_0}}{D}\sum_{i=1}^{n}\frac{(*)}{\alpha_i}$, where (*) is calculated below:

$$(*) \quad = \frac{W_{i1}}{\theta_0}\left[\theta_0(a_{11}b_{21}^{(i)}-b_{11}a_{21}^{(i)})+\theta_1(a_{11}b_{22}^{(i)}-b_{12}a_{22}^{(i)})+\theta_2(a_{11}b_{23}^{(i)}-b_{13}a_{21}^{(i)})\right]$$

$$+\ \frac{W_{i2}}{\theta_0}\left[\theta_1(a_{11}b_{21}^{(i)}-b_{11}a_{21}^{(i)})+\theta_0(a_{11}b_{22}^{(i)}-b_{12}a_{22}^{(i)})+\theta_3(a_{11}b_{23}^{(i)}-b_{23}a_{21}^{(i)})\right]$$

$$+\frac{W_{i3}}{\theta_0}\left[\theta_2(a_{11}b_{21}^{(i)}-b_{11}a_{21}^{(i)})+\theta_3(a_{11}b_{22}^{(i)}-b_{12}a_{22}^{(i)})+\theta_0(a_{11}b_{23}^{(i)}-b_{13}a_{21}^{(i)})\right]$$

$$=\frac{W_{i1}}{\theta_0}[-|\Sigma|\Delta(2b_{13}+b_{12})]+\frac{W_{i2}}{\theta_0}[-|\Sigma|\Delta(b_{13}-b_{11})]+\frac{W_{i3}}{\theta_0}[|\Sigma|\Delta(2b_{11}+b_{12})]$$

$$=\frac{|\Sigma|\Delta}{\theta_0}[-(2b_{13}+b_{12})W_{i1}-(b_{13}-b_{11})W_{i2}+(2b_{11}+b_{12})W_{i3}],$$

where W_{i1}, W_{i2}, and W_{i3} are given below.

Therefore, (2.9) becomes

$E(\hat{\beta}_1|Y_1 > C_1,\ Y_2 > C_2,\ Y_3 > C_3)$

$= \beta_1 + bias,$ (2.10)

where bias $= \frac{|\Sigma|\Delta}{D\sqrt{\theta_0}}\sum_{i=1}^{n}\frac{1}{\alpha_i}[-(2b_{13}+b_{12})W_{i1}-(b_{13}-b_{11})W_{i2}+(2b_{11}+b_{12})W_{i3}]$,

$D = n\Delta^2[a_{11}(2\theta_0^2-\theta_1^2-\theta_3^2-2\theta_1\theta_3+2\theta_0\theta_2)-(b_{11}-b_{13})^2]$,

$$a_{ij}=\frac{C_j-(\beta_0+\beta_1 t_{ij}+\tau_i)}{\sqrt{\theta_0}},\ j=1,2,3,$$

$$X_{ij}=\frac{Y_{ij}-(\beta_0+\beta_1 t_{ij}+\tau_i)}{\sqrt{\theta_0}},\ j=1,2,3,$$

$$A_{i,12}=\frac{\theta_0 a_{i2}-\theta_1 a_{i1}}{\sqrt{\theta_0^2-\theta_1^2}},\quad A_{i,21}=\frac{\theta_0 a_{i1}-\theta_1 a_{i2}}{\sqrt{\theta_0^2-\theta_1^2}},$$

$$A_{i,13}=\frac{\theta_0 a_{i3}-\theta_2 a_{i1}}{\sqrt{\theta_0^2-\theta_2^2}},\quad A_{i,31}=\frac{\theta_0 a_{i1}-\theta_2 a_{i3}}{\sqrt{\theta_0^2-\theta_2^2}},$$

$$A_{i,23}=\frac{\theta_0 a_{i3}-\theta_3 a_{i2}}{\sqrt{\theta_0^2-\theta_3^2}},\quad A_{i,32}=\frac{\theta_0 a_{i2}-\theta_3 a_{i3}}{\sqrt{\theta_0^2-\theta_3^2}},$$

$$W_{i1}=\phi(a_{i1})\,\Phi_2(A_{i,12},A_{i,13};\rho_{23.1}),\quad \rho_{23.1}=\frac{\theta_0\theta_3-\theta_1\theta_2}{\sqrt{(\theta_0^2-\theta_1^2)(\theta_0^2-\theta_2^2)}},$$

$$W_{i2}=\phi(a_{i2})\,\Phi_2(A_{i,21},A_{i,23};\rho_{13.2}),\quad \rho_{13.2}=\frac{\theta_0\theta_2-\theta_1\theta_3}{\sqrt{(\theta_0^2-\theta_1^2)(\theta_0^2-\theta_3^2)}},$$

$$W_{i3}=\phi(a_{i3})\,\Phi_2(A_{i,31},A_{i,32};\rho_{12.3}),\quad \rho_{12.3}=\frac{\theta_0\theta_2-\theta_2\theta_2}{\sqrt{(\theta_0^2-\theta_2^2)(\theta_0^2-\theta_3^2)}},$$

$\alpha_i=P(X_{i1}>a_{i1},X_{i2}>a_{i2},X_{i3}>a_{i3})$, where

$[X_{i1}X_{i2}X_{i3}]'\sim N_3(0,\Sigma)$, $\Sigma=\begin{bmatrix}1 & \rho_{12.3} & \rho_{13.2}\\ \rho_{12.3} & 1 & \rho_{23.1}\\ \rho_{13.2} & \rho_{23.1} & 1\end{bmatrix}$, ϕ is the standard normal density and $\Phi_2(x,y,r)$ is the survival of a standard bivariate normal distribution with correlation r.

<u>Special cases</u>:

1. If we take $\Sigma=\begin{bmatrix}\theta_0 & \theta_1 & \theta_2\\ \theta_1 & \theta_0 & \theta_3\\ \theta_2 & \theta_3 & \theta_0\end{bmatrix}$ and truncation at C_2, then

   $a_{i1}\to-\infty$, $a_{i3}\to-\infty$.

   Here $W_{i1}=\phi(a_{i1})\,\Phi_2(A_{i,12},A_{i,13};\rho_{23.1})$

   $=(0)\,(finite)=0$

   $W_{i3}=\phi(a_{i3})\,\Phi_2(A_{i,31},A_{i,32};\rho_{12,3})$

   $=0$ (finite) $=0$

$$W_{i2}=\phi(a_{i2})\,\Phi_2(A_{i,21},A_{i,23};\rho_{13.2})$$

$$=\ \phi(a_{i2})\,1=\phi(a_{i2})$$

and $\alpha_i=P(X_{i2}>a_{i2})$.

$$\therefore E(\hat{\beta}_1|Y_2\ >\ C_2)=\beta_1-\frac{|\Sigma|\Delta}{D\sqrt{\theta_0}}(b_{13}-b_{11})\sum_{i=1}^{n}\frac{\phi(a_{i2})}{P(X_{i2}>a_{i2})}$$

$$=\ \beta_1-\frac{|\Sigma|\Delta}{D\sqrt{\theta_0}}(b_{13}-b_{11})\sum_{i=1}^{n}\ (\text{hazard of standard normal at } a_{i2}).$$

2. Let $a_{i1}=a_{i3}$, then $A_{i,13}=A_{i,31}$. If we also take $\theta_1=\theta_3$, then $A_{i,12}=A_{i,32}$, $\rho_{23.1}=\rho_{12.3}$ and $b_{11}=b_{13}$. In this case bias = 0.

   $\therefore E(\hat{\beta}_1|Y_1>C_1,\ Y_2>C_2,\ Y_3>C_3)=\beta_1$.

3. If we truncate at C_2 and let $C_1\to-\infty, C_3\to-\infty$, and take the banded case $(\theta_1=\theta_3)$, then bias = 0 and $\hat{\beta}_1$ becomes an unbiased estimator of β_1.

3. ANALYSIS OF RANCH HAND DATA SETS

In this section four tables are provided based on the Analyses of Ranch Hand Data Sets. In each table, the first column gives the cut off points for log dioxin, the second column lists the number of observations involved in the analysis, third and fourth columns give the results based on the use of statistical analysis system (SAS) and some of the later columns are based on theoretical results for $\hat{\beta}_1$, bias and variance of $\hat{\beta}_1$

(under truncation). Bias and variance of $\hat{\beta}_1$ are obtained by the use of Tallis (1961) and McGill (1992).

3.1 Analysis of 1982, 1987 Ranch Hand Data

Model: $Y_{ij}=\beta_0+\beta_1 t_{ij}+\tau_i+\epsilon_{ij}, j=1,2; i=1,2,\ldots,n,$ same as (2.3) except j = 1,2

$$\sum = \sigma^2\begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}.$$

Formula: $\hat{\beta}_1 = -\frac{1}{n\Delta}\sum_{i=1}^{n}(Y_{i1}-Y_{i2}),$

$$E(\hat{\beta}_1 | Y_{i1}>C_1, Y_{i2}>C_2) = \beta_1 - \frac{\sigma(1-\rho)}{n\Delta}\sum_{i=1}^{n}\frac{[\phi(a_{i1})\Phi(A_{i,12})-\phi(a_{i2})\Phi(A_{i,21})]}{\alpha_i}$$

$= \beta_1$ - Bias,

Bias $= \frac{\sigma(1-\rho)}{n\Delta}\sum_{i=1}^{n}[\phi(a_{i1})\Phi(A_{i,12})-\phi(a_{i2})\Phi(A_{i,21})]/\alpha_i,$ (Δ = 5 years).

σ^2 = 1.045349
σ = 1.022423
μ = 2.5731

calculated on the basis of 1987 unrestricted data on 889 subjects

$\rho_{12}=\rho_{82,87}$ = .88497 (for calculations, see Mee & Chua (1991)).

TABLE 1

TRUNCATION POINT	n	$\hat{\beta}_1$ (SAS)	Var $(\hat{\beta}_1)$ (SAS)
Y_{82}>log10, Y_{87}>log10	387	-.05232519	$(.00430247)^2$
Y_{82}>log10+5(.05794781) Y_{87}>log10	365	-.05794781	$(.00429569)^2$

Table 1 (continued)

		Formula				
				McGill Tallis		
TRUNCATION POINT	n	$\hat{\beta}_1$	Sample variance	Bias	Var($\hat{\beta}_1$)	MSE
Y_{82}>log10, Y_{87}>log10	387	-.05602	.008036	.0036093	.000023298	.000036325
Y_{82}>log10+ 5(.05794781) Y_{87}>log10	365	-.06203	.007561	.00017057	.000024415	.000024441

3.2 Analysis of 1987, 1992 Ranch Hand Data

Model: $Y_{ij} = \beta_o + \beta_1 t_{ij} + \tau_i + \epsilon_{ij},\ j=2,3, i=1,2,\ldots,n,$

$\Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\sigma^2$ (same as in section 3.1).

Formula: $\hat{\beta}_1 = -\dfrac{1}{n\Delta}\sum_{i=1}^{n}(Y_{i2}-Y_{i3}),$

$$E(\hat{\beta}_1|Y_{i2}>C_2, Y_{i3}>C_3) = \beta_1 - \frac{\sigma(1-\rho)}{n\Delta}\sum_{i=1}^{n}\frac{1}{\alpha_i}[\phi(a_{i2})\Phi(A_{i,23}) - \phi(a_{i3})\Phi(A_{i,32})]$$

$$= \beta_1 - \text{Bias},$$

$$\text{Bias} = \frac{\sigma(1-\rho)}{n\Delta}\sum_{i=1}^{n}\frac{1}{\alpha_i}[\phi(a_{i2})\Phi(A_{i,23}) - \phi(a_{i3})\Phi(A_{i,32})],\quad (\Delta = 5 \text{ years}).$$

$$\left.\begin{aligned} \sigma^2 &= 1.045349 \\ \sigma &= 1.022423 \\ \mu &= 2.5731 \end{aligned}\right\}\ \text{calculated on the basis of 1987 unrestricted data on 889 subjects}$$

$\rho = \rho_{87,92} = .82286$ (for calculations, see Mee & Chua (1991)).

TABLE II

TRUNCATION POINT	n	SAS $\hat{\beta}_1$	SAS $Var(\hat{\beta}_1)$
Y_{87}>log10, Y_{92}>log10	240	-.10302671	$(.00584632)^2$
Y_{87}>log10+5(.11103261) Y_{92}>log10	213	-.11103261	$(.00613092)^2$

Table II (continued)

TRUNCATION POINT	n	Formula $\hat{\beta}_1$	Formula Sample variance	Formula McGill Tallis Bias	Formula McGill Tallis $Var(\hat{\beta}_1)$	Formula MSE
Y_{87}>log10, Y_{92}>log10	240	-.10906	.009051	.01055	.0000562	.0001675
Y_{87}>log10 + 5(.11103261) Y_{92}>log10	213	-.11853	.008849	.0003629	.0000611	.00006123

3.3 Analysis of Ranch Hand Data (1982, 1992)

Model: $Y_{ij}=\beta_0+\beta_1 t_{ij}+\tau_i+\epsilon_{ij},\quad j=1,3; i=1,2,\ldots,n,$

$$\Sigma=\begin{bmatrix}1 & \rho\\ \rho & 1\end{bmatrix}\sigma^2 \text{ (same as in section 3.1).}$$

Formula: $\hat{\beta}_1=-\frac{1}{2n\Delta}\sum_{i=1}^{n}(Y_{i1}-Y_{i3}),$

$$E(\hat{\beta}_1|Y_{i1}>C_1, Y_{i3}>C_3) =$$

$$\beta_1-\frac{\sigma(1-\rho)}{2n\Delta}\sum_{i=1}^{n}[\phi(a_{i1})\Phi(A_{i,13})-\phi(a_{i3})\Phi(A_{i,31})]\frac{1}{\alpha_i}$$

$$=\beta_1-Bias,$$

$$\text{Bias} = \frac{\sigma(1-\rho)}{2n\Delta} \sum_{i=1}^{n} \frac{1}{\alpha_i} [\phi(a_{i1})\Phi(A_{i,13}) - \phi(a_{i3})\Phi(A_{i,31})]$$

$$\left.\begin{array}{l} \sigma^2 = 1.045349 \\ \sigma = 1.022423 \\ \mu = 2.5731 \end{array}\right\} \begin{array}{l} \text{calculated on the basis of} \\ \text{1987 unrestricted data} \\ \text{on 889 subjects} \end{array}$$

$$\rho = \rho_{82,92} = \frac{\rho_{82,87} + \rho_{87,92}}{2} = .853915.$$

TABLE III

TRUNCATION POINT	n	SAS $\hat{\beta}_1$	SAS $Var(\hat{\beta}_1)$
Y_{82}>log10, Y_{92}>log10	240	-.07607383	$(.00329018)^2$
Y_{82}>log10+10(.0953214) Y_{92}>log10	216	-.07953214	$(.00352296)^2$

Table III (continued)

			Formula			
				McGill Tallis		
TRUNCATION POINT	n	$\hat{\beta}_1$	Sample variance	Bias	$Var(\hat{\beta}_1)$	MSE
Y_{82}>log10, Y_{92}>log10	240	-.07982	.002792	.0055485	.00001203	.00004282
Y_{82}>log10 + 10(.07953214) Y_{92}>log10	216	-.08373	.002878	.00026988	.00001261	.00001268

3.4 Analysis of Ranch Hand Data (1982, 1987, 1992)

Model: $Y_{ij}=\beta_0+\beta_1 t_{ij}+\tau_i+\epsilon_{ij}, j=1,2,3; i=1,2,\ldots,n,$

$$\Sigma=\begin{bmatrix}1 & \rho & \rho^2\\ \rho & 1 & \rho\\ \rho^2 & \rho & 1\end{bmatrix}\sigma^2, \qquad \text{(same as (2.3)).}$$

$$\hat{\beta}_1=\frac{1}{D}\sum_{i=1}^{n}\{(a_{11}b_{21}^{(i)})-b_{11}a_{21}^{(i)})Y_{i1}+(a_{11}b_{22}^{(i)}-b_{12}a_{21}^{(i)})Y_{i2}+(a_{11}b_{23}^{(i)}-b_{13}a_{21}^{(i)})Y_{i3}\}$$

see (2.8).

$E(\hat{\beta}_1|Y_1>C_1, Y_2>c_2, Y_3>c_3)$

$$= \beta_1 - \frac{\Delta|\Sigma|}{\sqrt{\theta_0 D}}(2b_{11}+b_{12})\sum_{i=1}^{n}(-w_{i3}+w_{i1})/\alpha_i \quad \text{(see (2.10))}$$

$= \beta_1 -$ Bias,

$$\text{Bias}=\frac{\sigma(1-\rho^2)}{2n\Delta}\sum_{i=1}^{n}\frac{1}{\alpha_i}[\phi(a_{i1})\Phi_2(A_{i,12}A_{i,13};\frac{\rho}{\sqrt{1+\rho_2}})-\phi(a_{i3})\Phi_2(A_{i,31}A_{i,32};\frac{\rho}{\sqrt{1+\rho^2}})]$$

$\sigma^2 = 1.045349$ calculated on the basis of

$\sigma = 1.022423$ 1987 unrestricted data

$\mu = 2.5731$ on 889 subjects

$\rho=\frac{\rho_{82,87}+\rho_{87,92}}{2} = .853915$ (see sections 3.1 and 3.2).

TABLE IV

TRUNCATION POINT	n	$\hat{\beta}_1$ (SAS)	$Var(\hat{\beta}_1)$ (SAS)
$Y_{82}>\log 10, Y_{87}>\log 10, Y_{92}>\log 10$	240	-.07586201	$(.00398025)^2$
$Y_{82}>\log 10+(.07910304)(10)$ $Y_{87}>\log 10+5(.07910304)$ $Y_{92}>\log 10$	213	-.07910304	$(.0042136)^2$

Table IV (continued)

TRUNCATION POINT	n	$\hat{\beta}_1$	Sample variance (Formula)	Bias (Formula; McGill Tallis)	$Var(\hat{\beta}_1)$ (Formula; McGill Tallis)	MSE (Formula)
$Y_{82}>\log 10$, $Y_{87}>\log 10$ $Y_{92}>\log 10$	240	-.07982	.002792	.0012561	.00000369	.00003848
$Y_{82}>\log 10$ $+(.07910304)(10)$ $Y_{87}>\log 10+$ $5(.07910304)$ $Y_{92}>\log 10$	213	-.08365	.002912	.00004389	.00004424	.00004424

Inspection of (2.10) shows that the bias can be made equal to zero by truncating the data such that $a_{i1}=a_{i2}=a_{i3}$, or, equivalently, such that the observations lie above a straight line with slope $-\lambda$. This requires an iterative procedure, because λ is unknown. The first step is to truncate the data at log(10) at each time point and estimate λ using weighted least-squares with AR(1) correlation structure. The second step is to truncate the first measurement at $\log(10) + 2\hat{\lambda}\Delta$, the second at $\log(10) + \hat{\lambda}\Delta$ and the third at log(10) and estimate λ using

weighted least squares with the AR(1) covariance assumption. The third step is to repeat the process using the updated value of $\hat{\lambda}$. The process is repeated until the estimated value of λ is equal to the value used for truncation. The final estimate of λ is then unbiased. The same procedure is followed in all the other cases.

4. CONCLUSIONS AND RECOMMENDATIONS

In this project we show that, unless the data set is properly conditioned, ordinary weighted least squares estimates of λ are biased due to regression toward the mean. If the within-subject correlation matrix is banded, we show that the weighted least squares estimate of λ is unbiased when the data set is conditioned on all repeated measures being above a line with slope $-\lambda$. AR(1) and compound symmetry are special cases of the banded matrix. Therefore the same results hold for these two cases.

Finally, the analyses of the bias, using the Ranch Hand data sets with two measurements as well as three measurements are given. The analyses have been carried out only for AR(1) model. Considering the importance of this work, it is hoped that this work will be a significant contribution to the Air Force Health Study.

Notice that $\hat{\beta}$ has been obtained without considering the trunction factor while $E(\hat{\beta})$ and $\mathrm{Var}(\hat{\beta})$ have been calculated

under the truncated model. A future study, if funding is available, will consist of obtaining $\hat{\beta}$ under the truncated model. For this modified value of $\hat{\beta}$, $E(\hat{\beta})$ and $\mathrm{Var}(\hat{\beta})$ will be recalculated. It is hoped that half-life estimates, thus obtained, will be superior.

ACKNOWLEDGEMENTS

I would like to thank the Armstrong Laboratory and the Air Force Office of Scientific Research for sponsoring this research.

My special thanks are due to Dr. Joel E. Michalek for providing me some insight in to this problem. It was very nice to work with the scientists at the Armstrong Laboratory during summer months.

REFERENCES

James, K.E. (1973). Regression toward the mean in uncontrolled clinical studies. Biometrics, 29, 121-130.

McGill, J. I. (1992). The multivariate hazard gradient and moments o the truncated multinormal distribution. Communications in Statistics, Theory and Methods, 21(11), 3053-3060.

Mee, R. W. and Chua, T. C. (1991). Regression toward the Mean and the paired sample t-test. The American Statistician, 45(1), 39-42.

Senn, S.S. and Brown, R.A. (1985). Estimating treatment effects in clinical trials subject to regression to the mean. Biometrics, 41, 555-560.

Tallis, G.M. (1961). The moment-generating function of the truncated multinormal distribution. J. R. Statist. Soc. B, 23, 223-229.

Manfred Koch report unavailable at time of publication.

Mark Novotny report unavailable at time of publication.

Surface Fitting Three Dimensional Human Head Scan Data

Joseph H. Nurre
Assistant Professor
Department of Electrical Engineering & Computer Science

and

Max D. Corbin
Research Assistant
Department of Electrical Engineering & Computer Science

Ohio University
Athens, OH

Final Report for:
Summer Research Extension Program

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base
Washington, DC

and

Armstrong Laboratory

December 1995

SURFACE FITTING THREE DIMENSIONAL HUMAN HEAD SCAN DATA

Joseph H. Nurre
Assistant Professor
Department of Electrical Engineering & Computer Science
Ohio University

Max D. Corbin
Research Assistant
Department of Electrical Engineering & Computer Science
Ohio University

ABSTRACT

Full field surface data of cylindrically shaped objects, such as a human's head, can be quickly achieved by rotating a laser scanner and imaging system about the subject. B-spline surfaces can then be fitted to the measurements for data reduction and compatibility with NURBS based CAD systems. Techniques available for surface fit are subject to user input. Parameters, such as, the number of control points, tension (in the sense of a thin plate spline), etc. must be chosen to achieve optimal fit. This report will discuss software developed to fit uniform cubic B-spline surfaces to head scan data sets. Furthermore, the optimal choice of surface fitting parameters for such scan data sets is discussed. The software and techniques are to be used by the U.S. Air Force Computerized Anthropometric Research and Design Laboratory as part of their mission of providing a data repository for anthropometric information.

FITTING THREE DIMENSIONAL HUMAN HEAD SCAN DATA

Joseph H. Nurre and Max D. Corbin

1. INTRODUCTION

The Computerized Anthropometric Research and Design (CARD) Laboratory of the Human Engineering Division at Wright-Patterson Air Force Base, Ohio has employed a Cyberware ECHODigitizer, a laser scanning system well suited to capturing human head scans, to acquire surface data of the head and face. Currently, the CARD laboratory maintains a database of over 1000 scan images of human heads which include both military and civilian personnel. Specialized software to manipulate and display the head scan is a necessity. An interactive program called INTEGRATE has been created to handle this task with a wide variety of functions, implemented in a modular programming fashion.

Surface data recorded from a head scan is represented in cylindrical coordinates with a sampling resolution of 1.563mm in the latitudinal direction, and 0.7031 degrees of arc in the longitudinal direction. This results in a 256 by 512 array of radius values. To better utilize the head scan information, the data can be fitted to a surface function. Cubic B-spline surfaces have become a universally accepted modeling structure in computer graphics applications.[1] They represent one class of Non-Uniform Rational B-Spline Surfaces (NURBS).

Fitting the head scan data to B-spline surfaces would provide a number of benefits. First, the head scan can be represented more efficiently, with fewer data points. Large numbers of data points, located on smooth surfaces of the head, can be fitted with a single B-spline surface patch that requires only 16 control points for representation. A second benefit is the ability to manipulate scans quickly. B-splines, being of an affine invariant operation, can be translated and rotated by simply moving the control points. The

surfaces are continuous, and point by point transformations are unnecessary. Finally, the third benefit is portability to commercially available Computer Aided Engineering (CAE) packages. B-splines are a standard surface representation in CAE software used in designing appliances with human interfaces.

2. REVIEW OF CUBIC SPLINES

The surface function to be fitted to the head scan data is the two parameter cubic spline. To simplify the discussion, a one parameter cubic spline will first be considered. The one parameter cubic spline is a third degree polynomial that describes the path of a particle through time and space. The trace of this particle is the desired curve.

The particle path and hence a two dimensional curve can be represented by two polynomials:

$$x(t) = K_{0x} + K_{1x}t + K_{2x}t^2 + K_{3x}t^3 \tag{1}$$

$$y(t) = K_{0y} + K_{1y}t + K_{2y}t^2 + K_{3y}t^3 \tag{2}$$

where $0 \leq t \leq 1.0$

Equations (1) and (2) can be written as:

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} K_{0x} \\ K_{0y} \end{bmatrix} + \begin{bmatrix} K_{1x} \\ K_{1y} \end{bmatrix} t + \begin{bmatrix} K_{2x} \\ K_{2y} \end{bmatrix} t^2 + \begin{bmatrix} K_{3x} \\ K_{3y} \end{bmatrix} t^3 \tag{3}$$

Equation (3) represents four points in two dimensional space that are operated on to give a two dimensional curve. The parameter t and its coefficient points can be rewritten as:

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} B_{0x} \\ B_{0y} \end{bmatrix}(1-t)^3 + \begin{bmatrix} B_{1x} \\ B_{1y} \end{bmatrix}3t(1-t)^2 + \begin{bmatrix} B_{2x} \\ B_{2y} \end{bmatrix}3t^2(1-t) + \begin{bmatrix} B_{3x} \\ B_{3y} \end{bmatrix}t^3 \qquad (4)$$

where equations (3) and (4) can be made equivalent for the appropriate coefficient values.

Equation (4) is defined as the Bezier curve and is controlled by points B_0 B_1 B_2 B_3. The curve always interpolates points B_0, B_3 and is tangent to vectors $\overline{B_0B_1}$ and $\overline{B_2B_3}$ as shown in Figure 1.

A different set of functions can be used that define a curve called the uniform B-spline:

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \frac{1}{6}\left(\begin{vmatrix} C_{0x} \\ C_{0y} \end{vmatrix}(1-t)^3 + \begin{vmatrix} C_{1x} \\ C_{1y} \end{vmatrix}(3t^3-6t^2+4) + \begin{vmatrix} C_{2x} \\ C_{2y} \end{vmatrix}(-3t^3+3t^2+3t+1) + \begin{vmatrix} C_{3x} \\ C_{3y} \end{vmatrix}t^3\right) \quad (5)$$

To understand the B-spline, one can find Bezier control points from the control points in the equation above.[2] First, draw a line between C_0-C_1, C_1-C_2 and C_2-C_3. These lines are then divided into three equal parts. Two additional lines are drawn from points on these divided lines as shown in Figure 2. Bezier control points B_0 and B_3 fall on the respective bisectors of these two new lines. Control points B_1 and B_2 are located on the hash marks of the line between C_1-C_2.

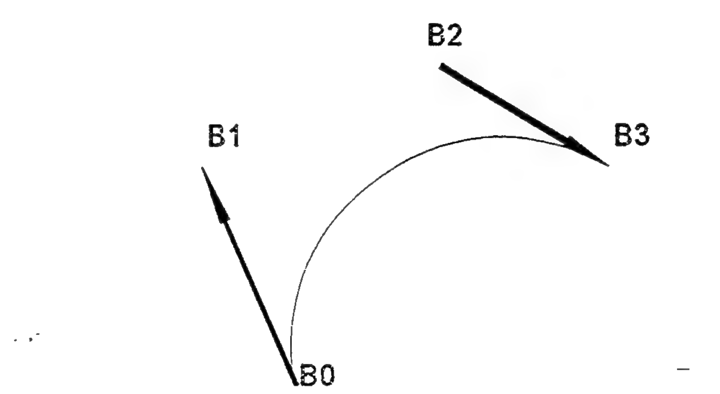


Figure 1. The influence of control points on a Bezier curve.

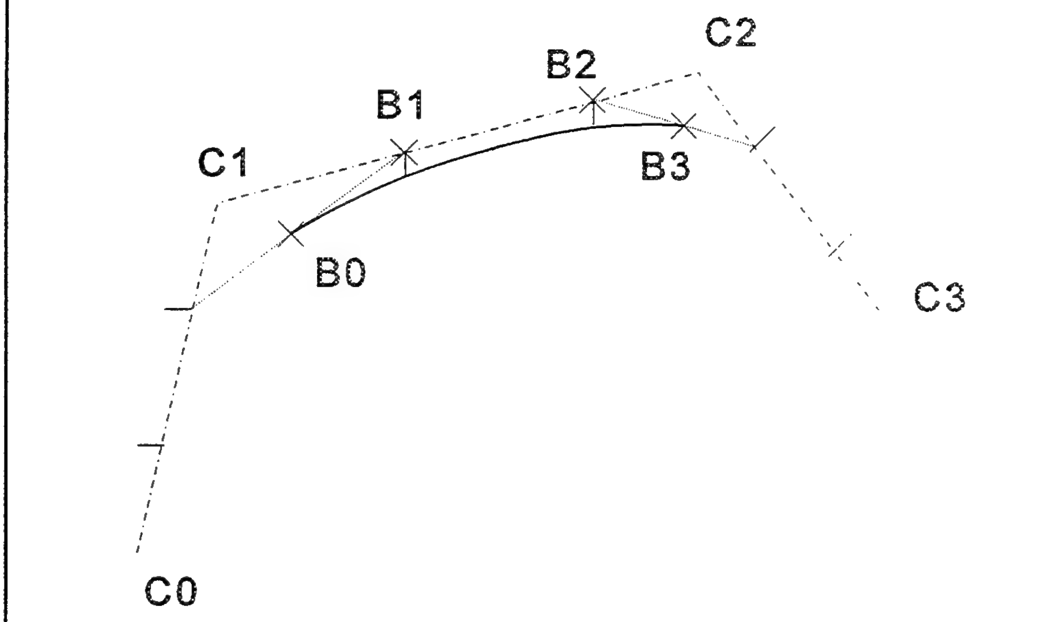


Figure 2. A Bezier curve can be obtained from the control points of a uniform B-spline as shown.

Control points of multiple B-splines always overlap. In Figure 3, two cubic B-splines are controlled by the points C_0 C_1 C_2 C_3 C_4. The first spline uses C_0 C_1 C_2 C_3 while the second spline is controlled by C_1 C_2 C_3 C_4. If one derives the Bezier control points for each of these splines, it is clear that the B_3 control point of the first spline is equal to B_0 of the second spline. This is known as C^0 continuity. In other words, the two splines always touch. Furthermore, the $\overline{B_2B_3}$ vector of the first spline is equal to the $\overline{B_0B_1}$ vector of the second spline. This is known as C^1 continuity. The tangent of the two splines is identical where they meet. It can be shown that in fact B-splines have C^2 continuity where they meet. The polynomials of the splines must be differentiated at least three times for a discontinuity between the functions to appear.

A two parameter cubic spline must be used to describe a surface. Analysis of the surface spline proceeds in an analogous fashion to the curve,

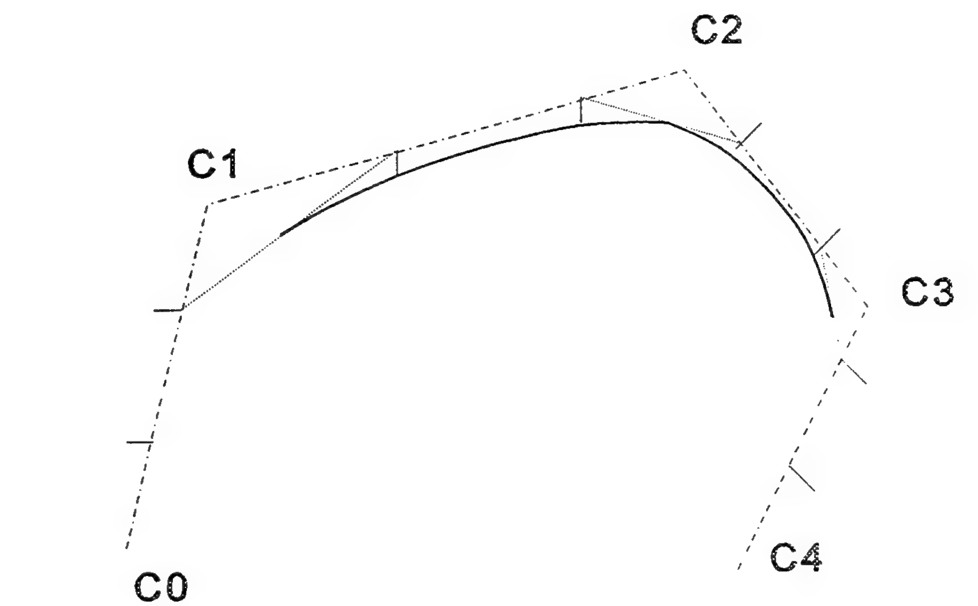


Figure 3. Two uniform B-splines exhibit C^2 continuity.

with the resulting functions once again demonstrating C^2 continuity.

3. FITTING DATA TO SPLINES

Splines can be fitted to data in a least squares sense similar to the method used to fit a straight line function to data.[3,4] In linear least squares regression, the function:

$$w = f(t) = k_1 t + k_0 \tag{6}$$

is to be fitted to data where t_i is a known ordinate and w_i is the measured response. The value of k's that best generate a line which fit the data in a least squares sense, will minimize the functional:

$$\tau = \sum_{i=0}^{N-1} (w_i - (k_1 t_i + k_0))^2 \tag{7}$$

From calculus, the minimum occurs when:

$$\frac{\partial \tau}{\partial k_j} = 0 \qquad j = 0,1 \tag{8}$$

Expressing equation (7) term by term gives:

$$\begin{aligned}\tau = \ & \{w_0 - (k_1 t_0 + k_0)\}^2 + \\ & \{w_1 - (k_1 t_1 + k_0)\}^2 + \\ & \{w_2 - (k_1 t_2 + k_0)\}^2 + \\ & \cdots\cdots\cdots\cdots\end{aligned}$$

which gives the pseudo matrix form:

$$\tau = \sum \left(\begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ \bullet \\ \bullet \end{bmatrix} - \begin{bmatrix} t_0 & 1 \\ t_1 & 1 \\ t_2 & 1 \\ \bullet & \\ \bullet & \end{bmatrix} \begin{bmatrix} k_1 \\ k_0 \end{bmatrix} \right)^2 \tag{9}$$

When equation (9) is differentiated with respect to the constants, k_j, and set to zero, we get:

$$0 = 2 \begin{bmatrix} t_0 & t_1 & t_2 & \bullet & \bullet \\ 1 & 1 & 1 & \bullet & \bullet \end{bmatrix} \left(\begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ \bullet \\ \bullet \end{bmatrix} - \begin{bmatrix} t_0 & 1 \\ t_1 & 1 \\ t_2 & 1 \\ \bullet & \\ \bullet & \end{bmatrix} \begin{bmatrix} k_1 \\ k_0 \end{bmatrix} \right) \tag{10}$$

This gives,

$$\begin{bmatrix} \sum w_i t_i \\ \sum w_i \end{bmatrix} = \begin{bmatrix} \sum t_i^2 & \sum t_i \\ \sum t_i & N \end{bmatrix} \begin{bmatrix} k_1 \\ k_0 \end{bmatrix} \tag{11}$$

or simply:

$$\mathbf{b}_L = \mathbf{A}_L \mathbf{z}_L \tag{12}$$

When equation (12) is non-singular, the unknown $\mathbf{z}_L$ can be solved for using Gaussian elimination or other appropriate methods.[5] An equation, similar to equation (12), can be derived for spline functions. The vector of constants k_j becomes a vector of spline control points. Data points and appropriate functions of t can be used to construct an equation similar to equation (9). Proceeding in a similar manner as given above, we get a matrix equation:

$$\mathbf{A}_B \mathbf{z} = \mathbf{b} \tag{13}$$

Equation (13) is non-singular and can be solved for the B-spline control points. However, the resulting spline tends to fluctuate due to the noise in the measured data. To obtain a smooth curve, an additional "regularization" term is added to equation (13) to further constrain the spline.

Regularization is a theory developed early this century which has been used to solve ill-posed problems. An ill-posed problem may have multiple solutions. Regularization restricts the class of admissible solutions, creating a well-posed problem, by introducing a "stabilizing function".[6] An integral part of the regularization process is a parameter which controls the

tradeoff between the "closeness" of the solution to the data and its degree of "smoothness" as measured by the stabilizing function. The mathematical expression of regularization in one dimension is given as follows:

$$E = \int_{\Omega}(w-f)^2 dx + \lambda\int_{\Omega} S(f)^2 dx \tag{14}$$

In this functional, the first term is the continuous least squares measure of the closeness of the solution function f to the data measurement function w. The second term is considered to be the stabilizer and assures the smoothness of f.

For surface data, a particularly useful stabilizer is to minimize the first and second partial derivative of the function.[7,8] The first partial derivative gives the surface elastic properties which become taut with increased emphasis. The second partial derivative causes the surface to act like a thin sheet of metal. Equation (14) for a one parameter spline with first and second derivative stabilizers is given as:

$$\tau = \sum_{i=0}^{N-1}(w_i - f(t_i))^2 + \lambda\int(\alpha f'(u) + \beta f''(u))^2 du \tag{15}$$

The integral in equation (15) can be solved exactly due to the underlying simplicity of a polynomial. Minimizing just the integral part of equation (15) with respect to the control points gives a matrix:

$$A_2 z \tag{16}$$

Finding the minimum of equation (15) is simply a matter of finding the solution to the matrix equation given below using standard numerical techniques:

$$(A_B + A_2)z = b \tag{17}$$

The discussion above is generalized to a two parameter spline, useful for fitting surfaces.

4. SURFACER

Surface fitting is based on minimizing the equation:

$$\tau = \sum_{i=0}^{N-1} (w_i - f(u_i, v_i))^2 + \lambda \iint \rho (f_{uu}^2 + 2f_{uv}^2 + f_{vv}^2) + (1-\rho)(f_u^2 + f_v^2)\, du dv \tag{18}$$

which is an extension of equation (15) to the two parameter case. The regularization term is based on a thin plate model. The parameter ρ determines how much of the smoothing(λ) is based on 1st or 2nd derivative information in the spline.

A two parameter cubic spline fitting surface program was written that allows for variation in the number of control points and regularization term values. The program is configured to do a surface fit on head scan or other data sets that are represented on a rectangular grid wrapped around a single axis. In other words, the data is cylindrical in nature. The control points of the surface to be fitted are placed in a rectangular array with rows representing the height of the data set (Z axis) and where columns represent the unwrapped radial angle (θ, R axis).

The algorithm used is based on non-rational uniform B-splines which are a subset of NURBS. Requiring the knots to be uniform allows for maximum data compression potential since only the spline control points need to be stored and not accompanying knot vectors. The use of uniform splines also increases the number of possible applications that would be able to use the resulting surface by making direct Bezier surface conversion possible, and allows for important coding simplifications in the equations.

4.1 Software Overview

The program structure is based around three main elements shown in Figure 4: a generic surfacer core, a user interface, and file I/O. The code was designed so that the surfacer core can be interfaced to other program code (such as CARD lab's INTEGRATE software) with minimal changes. Parameters and scan data are passed to the surfacer through a defined data structure. The surfacer then returns this structure with control points and surface error computations. The user interface handles user queries, parameter settings, window management, and the display of generated surfaces. The file I/O both manages loading/saving the raw data as well as importing/exporting other data formats.

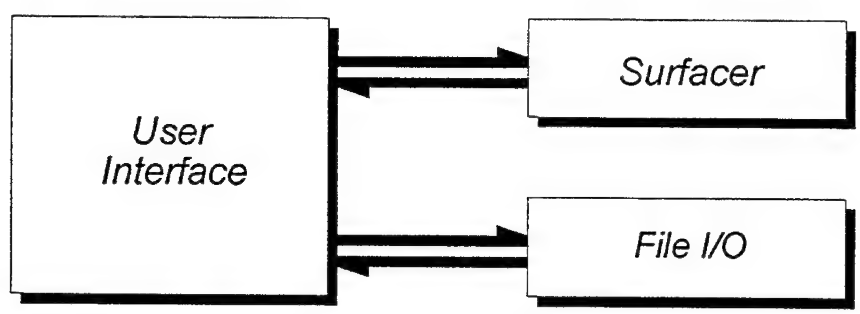


Figure 4. Block diagram of surfacer program.

4.2 Surfacer Core

The surfacer core performs the math necessary to fit B-spline surfaces to the head scan data. All surface fitting algorithms suffer from memory requirements that grow exponentially as the size of the surface fit is increased. An example of this would be a 512x136 surface fit. The $(A_B + A_2)$ matrix would be of size (512x136)x(512x136) and consist of 4.8486×10^9 cells. With double point precision in calculations, each cell requires 8 bytes of memory. This results in 38.8 Gigabytes of RAM required for storage. Other matrices used in the surfacing calculations are negligible in size compared to this one.

Two issues must be addressed to solve the size problem. First, the information in the $(A_B + A_2)$ matrix must be saved in a memory efficient fashion. Secondly, a method to find the inverse of $(A_B + A_2)$ from its memory

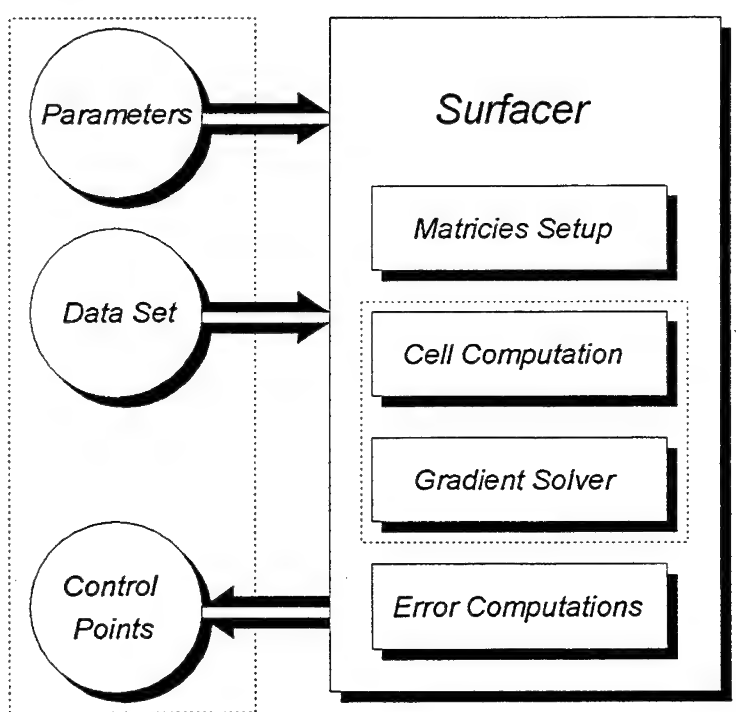


Figure 5. Block diagram of the surfacer core.

efficient representation must be designed. The surfacer core reduces memory requirements by taking advantage of properties of uniform B-splines, and properties of the regularization term. For example approximately 90% of the $(\mathbf{A}_B + \mathbf{A}_2)$ matrix values are zeros. A great deal of symmetry is also present in the matrix due to the symmetry of equation (5). The $(\mathbf{A}_B + \mathbf{A}_2)$ matrix can be broken into six smaller core matrices.[3,8,9] Any element in $(\mathbf{A}_B + \mathbf{A}_2)$ is found from computations on the core matrices without storing the actual large matrix. In the case of the 512x136 fit, three matrices of size 512x512 and three of size 136x136 are required. The same matrix that required Gigabytes of RAM can be stored in $3*8*(512^2+136^2)=6.74$ Megabytes.

Standard numerical techniques fit the surface equations by finding the inverse of the matrix $(\mathbf{A}_B + \mathbf{A}_2)$. Methods such as Gauss-Jordan elimination and LU decomposition give solutions exact to within the numerical representation of a double precision number. The inverse calculations are done "in place" and require that $(\mathbf{A}_B + \mathbf{A}_2)$ be stored in its entirety. The use of a gradient solver introduces some error into the solutions, but only requires a few small matrices in addition to the core matrices during calculations. The tolerance of the error, fortunately, can be specified. By applying a gradient solver, surface fits where the number of control points equals the number of data points (typically 70000 or more) can be accomplished in under 20 Megabytes of RAM. The use of a gradient solver also reduces the amount of time required for computations. Results can be had in minutes instead of hours or even days.

4.3 User Interface

The user environment or interface shown in Figure 6 gives an interactive method for controlling the parameters being passed to the surfacer. It also controls file operations. The interface is built around a collection of "widgets" or tools that make an interactive dialog with the user easier from a

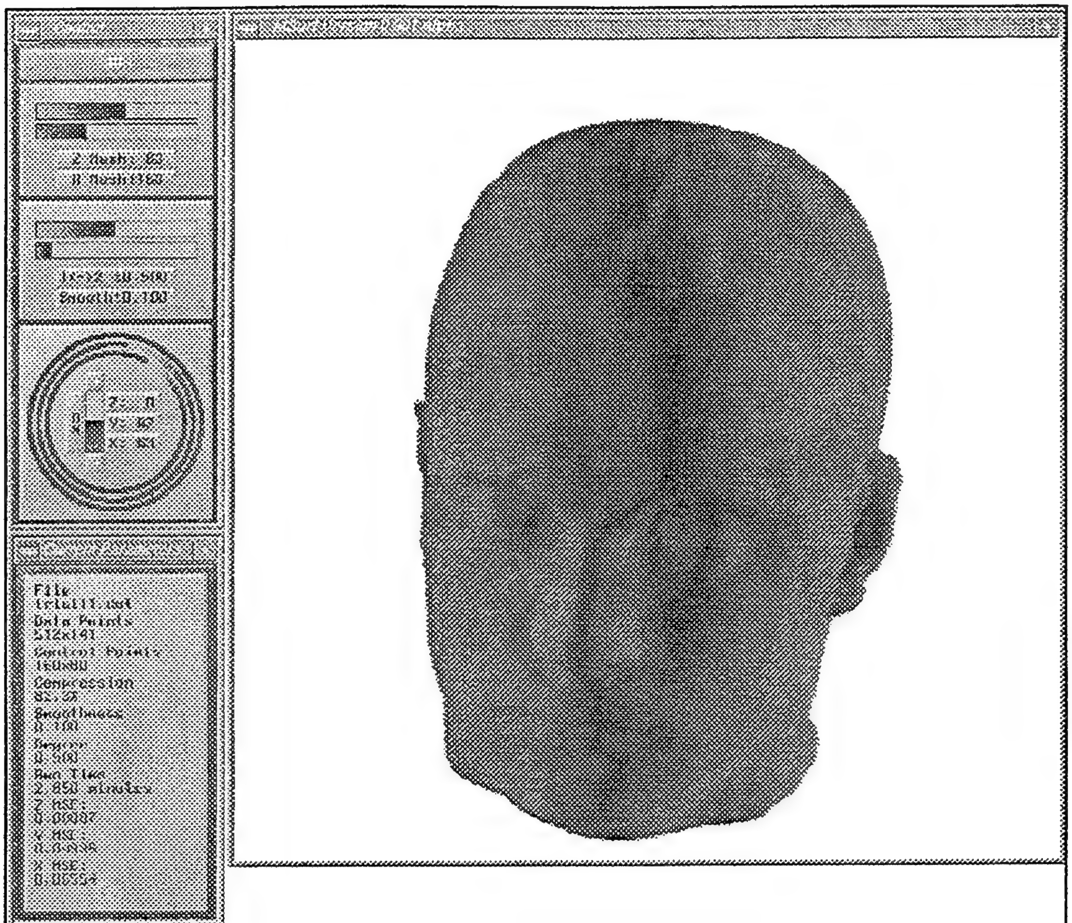


Figure 6. The surfacer's user environment consists of the *surface display* (large window), *current settings* (window lower left), and the *control panel* (window upper left). Interactive menus are available by clicking on the *surface display*.

programmers viewpoint. These widgets include: menus, information queries, progress bars, information display, buttons, and error message handling. These are used to build the *surface display*, *interactive menus*, *current settings*, and *control panel*.

The *surface display* window shows the loaded data sets as a collection of small cross hairs located at each data point in space. If a surface has been computed then it is shown as a shaded NURBS surface. Data points and a fitted surface can be seen centered in the display windows in Figures 7 and 8 respectively.

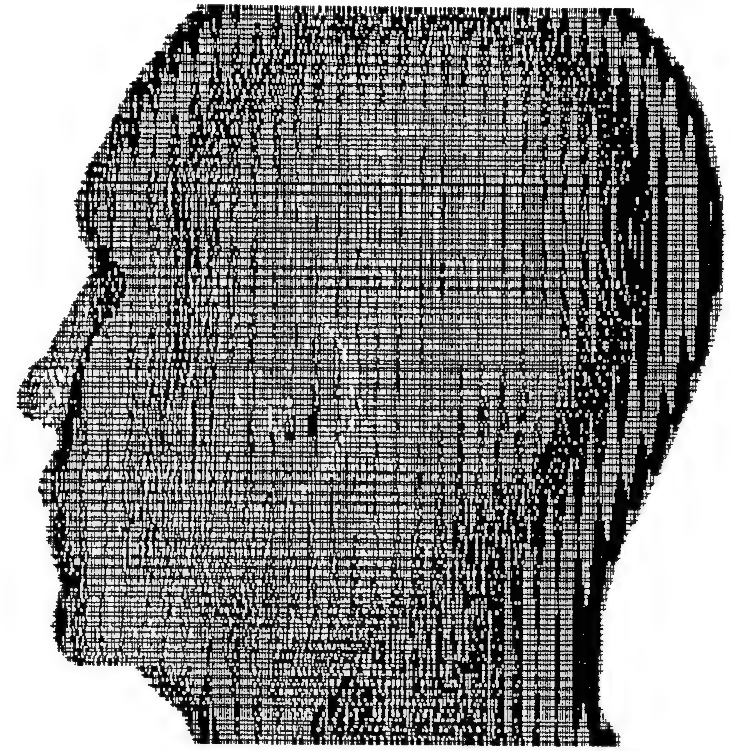


Figure 7. Point representation of a head scan data set.

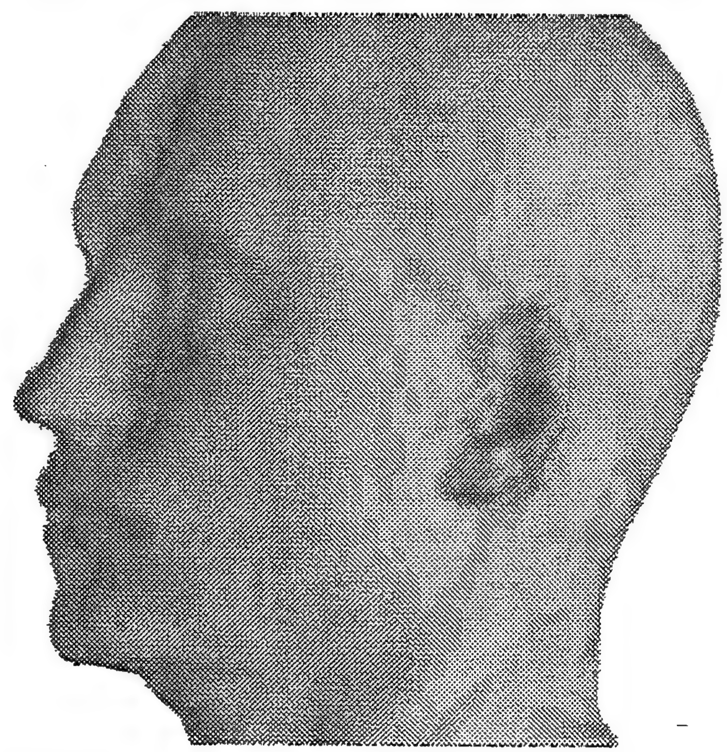


Figure 8. A surface fitted over the data points of Figure 7.

Interactive menus are available from the display window. These menus control viewing angles, file I/O, surface display, and surfacer parameters.

The *control panel* shown in Figure 9 gives a graphical interface to the menus of the display window. Viewing controls and surfacer parameters can be quickly and interactively set from this tool. The controls allow the viewing angle and viewing scale to be set by grabbing the appropriate bar with the mouse. The viewing angles are controlled with the three buttons of the mouse corresponding to the X, Y, and Z arcs located at the bottom of the window. On slower machines, surface redraw after a view change can be very time consuming. An ON/OFF switch, within the X, Y, and Z arcs exists for turning off surface redraw so viewing angles can be quickly set. Viewing scale is set with the bar next to the ON/OFF switch. The surfacer parameters are adjusted in the top of the window by grabbing a bar and sliding it to change the values, or clicking on a button to manually enter values. The RUN button on top executes the surfacer with the user specified parameters.

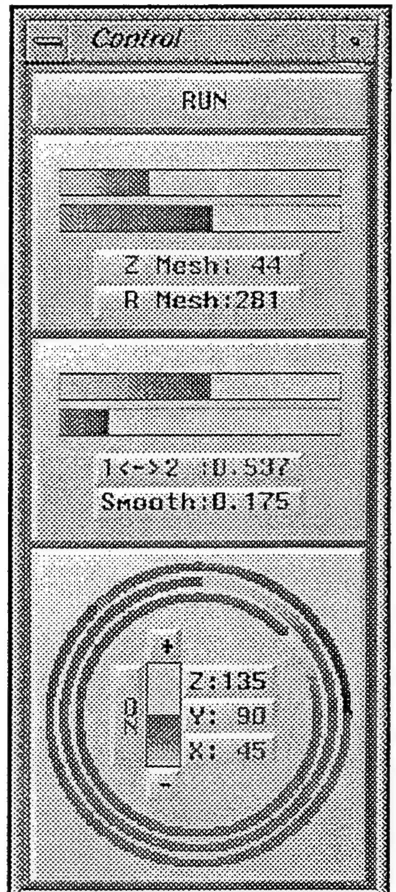


Figure 9.
Control Panel.

The *current settings* window seen in Figure 10 shows important parameter data such as the name of the loaded data set, size of the loaded data set, size of the fitted surface, and surface mean square errors. The errors are computed for an XYZ coordinate system. Information about computation time and compression potentials are also present in this window.

It should be noted that command line operation of the code is also

available. This allows the user to run the surfacer in the background and save results. The command line mode was used extensively in a batch file to generate the results in section 5.

4.4 File I/O

The file I/O module is over viewed in Figure 11 and handles head scan data retrieval and saving. Currently four functions are supported: import ASCII data set, load native data set, export IGES surface, and save native data set. Head scan data import reads ASCII files generated by CARD lab's INTEGRATE software. The file I/O module can also read and write a native ASCII format specific to the surfacer that contains a head scan, a fitted surface if available, and any parameters used to generate the surface. Finally, the file I/O module has the ability to write an industry standard IGES file for fitted surfaces. This utility allows fitted surfaces to be manipulated by other CAD programs. It should be noted that importing the INTEGRATE ASCII format and exporting an IGES surface are

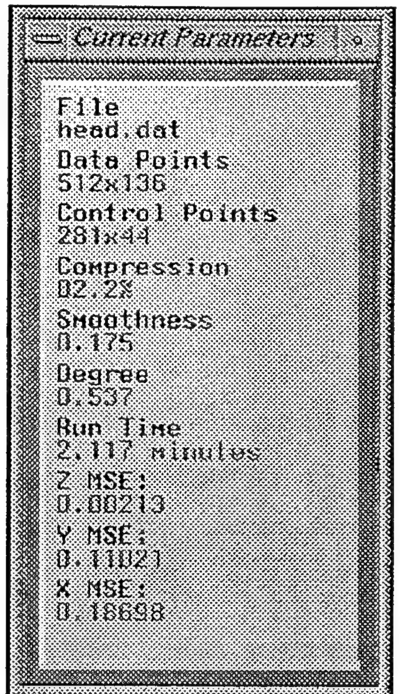


Figure 10.
Current Settings.

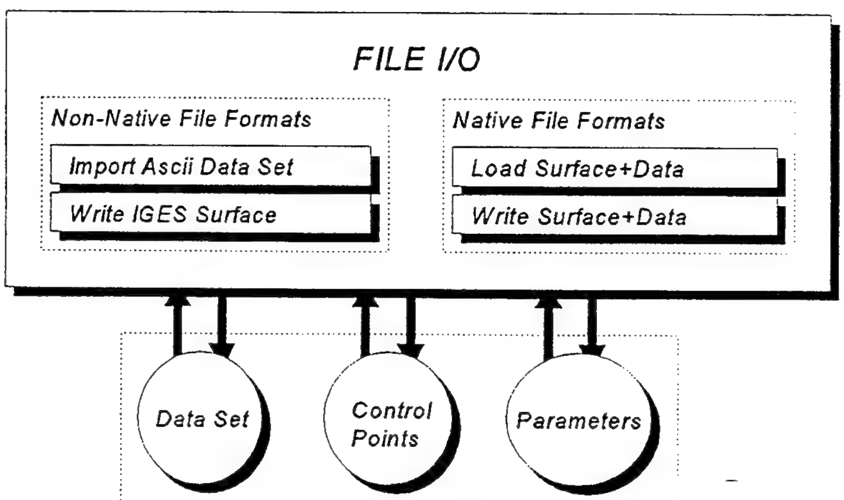


Figure 11. Block diagram of file operations.

the only file operations available from the command line mode of the program.

5. ACHIEVING OPTIMUM FIT

Figure 12 shows a typical Cyberware head scan data set of mesh size 512x136. Figure 13 shows this set fitted with surfaces of size 128x32, 281x74, and 512x136 respectively by the surfacer software. The figures show that a great potential for data compression exists using surfacing data. Choosing the fitting parameters correctly, however, is important for maintaining the integrity of the underlying anthropometric information.

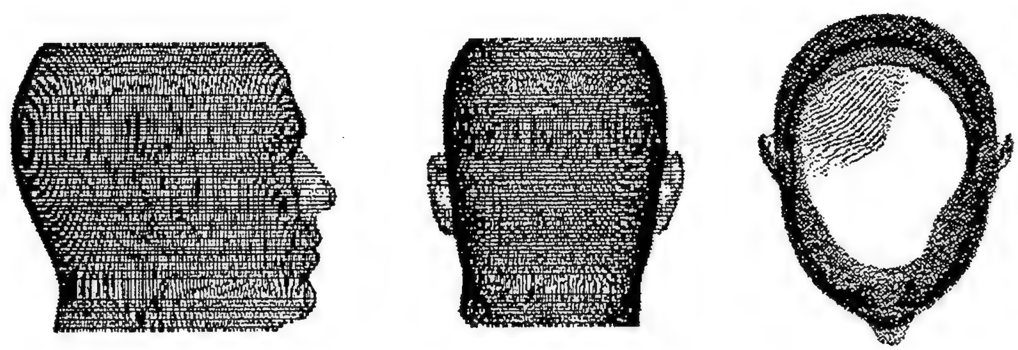


Figure 12. Various views of a head scan data set.

The optimization of surface fit parameters will be explored using data collected from thirty head scans.[10] Optimum fit will be defined as minimum mean square error (MSE) between the surface and data points. The mean square is computed as:

$$MSE = \frac{\sum_{i=0}^{N-1} (w_i - f(u_i, v_i))^2}{N} \tag{19}$$

The first thing to consider in choosing parameters is the size of the control point mesh needed to achieve a good fit. Figure 14 shows how errors

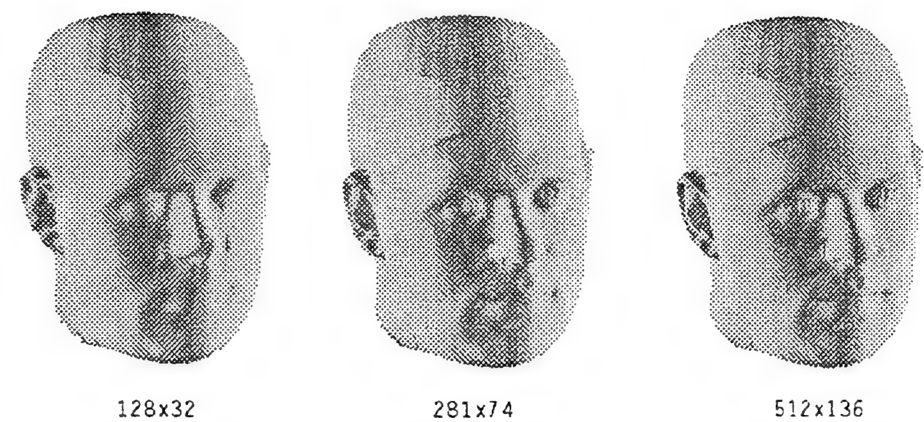


Figure 13. Different surface fits applied to the same data set.

along the X, Y, and Z axes change as the total points fitted change. As the number of fitted points increase the errors decrease. The graph shows that fit error is relatively small for surprisingly small surface meshes. As predicted by the graph, and as seen in Figure 13, a 281x74 (~20000 points)

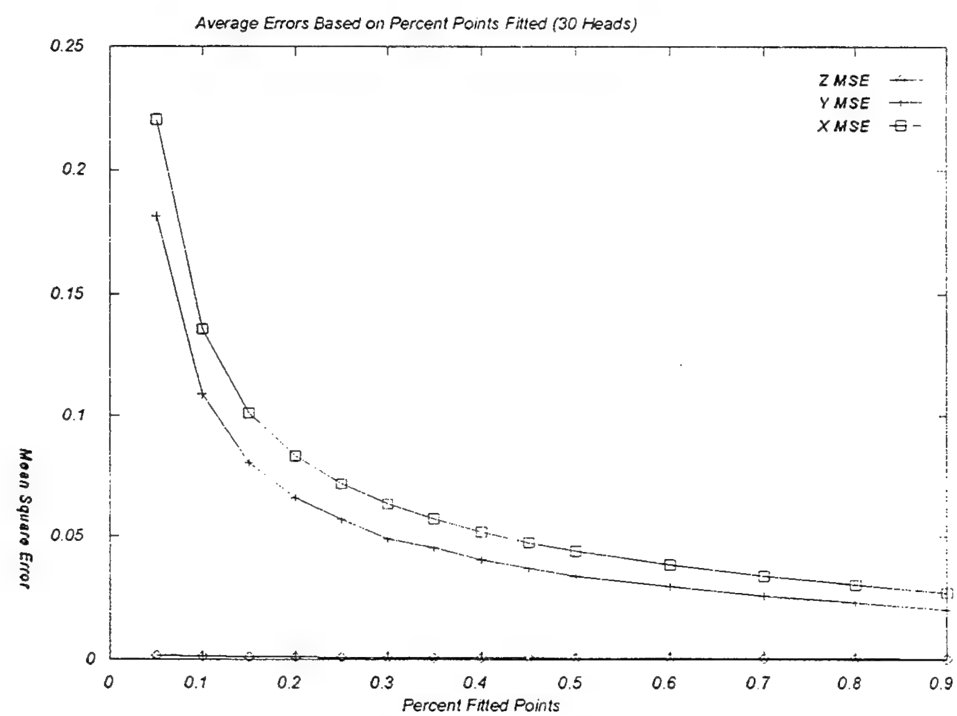


Figure 14. Average MSE errors computed for 30 heads.

mesh does an excellent job of reconstructing a surface for the head data in Figure 12. Furthermore this surface requires only 30% of the storage needed for the original data set. The graph suggests that mean square error does not significantly rise for fits as small as 14%. In practice, a fit this small shows loss of surface discontinuities (i.e. areas around the nose in a head scan) due to the smoothing properties of splines.

In addition to the total number of control points, the ratio of rows to columns is also important. A 512x5 surface has more error than a 64x40 surface yet requires the same storage. Figure 15 shows the error for the relation of Radial size(R) to the Height(Z). This graph shows that a minimum condition for surface fitting given a fixed fitting size (i.e. 30%), occurs when the number of control points along the radial axis is 1.6 times the number along

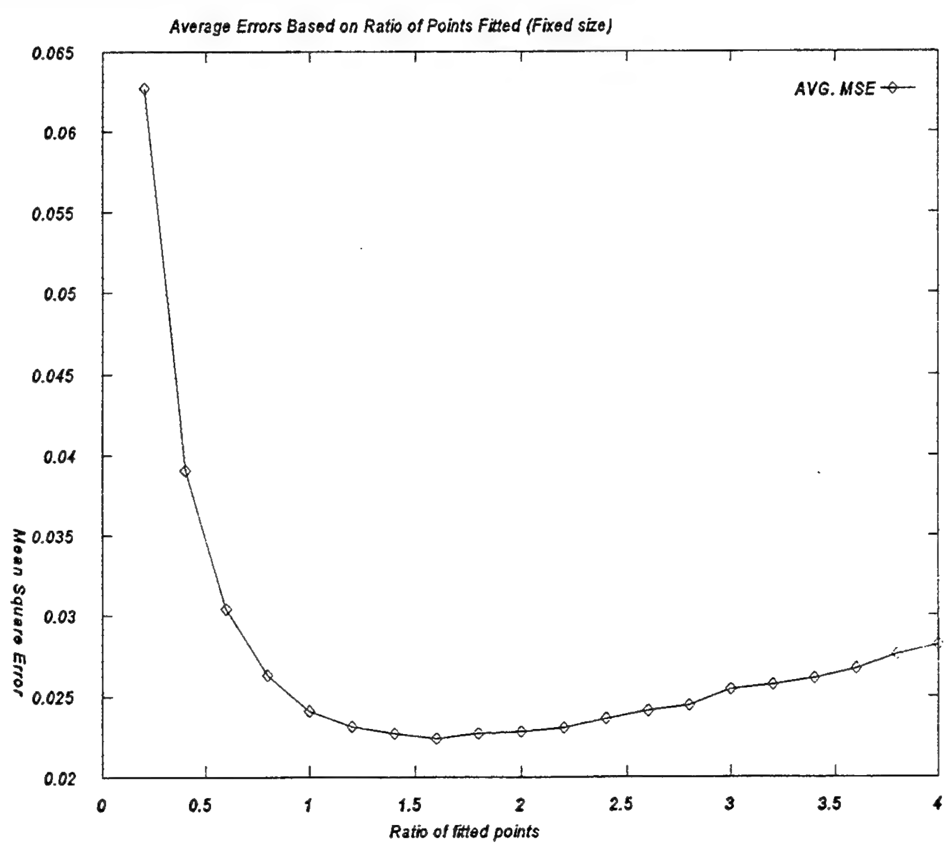


Figure 15. Average MSE error based on ratio of surface fit.

the height axis (R/Z = 1.6).

Regularization parameters λ and ρ will also affect fit. Mean square error gives little insight into choosing these parameters. Reducing surface smoothness by decreasing the size of the regularization parameter, λ, causes the fitted splines to oscillate in an effort to fit the splines closer to respective data points. Minimum error in a MSE sense is achieved, but severe ripples result near surface discontinuities. A value should be chosen that is as small as possible without causing these oscillations. For the trial head scan data sets, a value of λ=0.1 was used with good results. The parameter ρ also affects system error. Practice suggests that less error can be achieved by biasing the fit toward 1st degree information. However, if ρ is too small, surface oscillations will again appear. For the collection of data sets investigated, $\rho \geq 0.25$ prevents ripples.

6. CONCLUSIONS

A software system for fitting surfaces to three dimensional head scan data has been presented. The software has a modular design with a surfacer core, user interface and file I/O. The surfacer core uses an innovative data representation and solver to avoid the problem of large data sets. This advancement allows the software to run on smaller and cheaper workstations than otherwise would have been required. The user interface and file I/o modules presented a simple, user friendly environment for fitting surface data. The modular design also allows for easy integration of the code with the INTEGRATE software currently available at the CARD lab. The ability to export surfaces as IGES files enables users to utilize the head scan information in other CAD/CAM based applications.

Experiments were conducted to determine parameters which would achieve an optimal fit in the mean square sense, for the head scan data. First, the total number of control points fitted to the surface should be around 30% or more of the number of data points. Depending on how much surface detail can

be lost, this number could go as low as 14%. Secondly, the number of control points fitted around the radial axis should be 1.6 times the number of points fitted along the height axis. Finally, ρ and λ should be as small as possible and still large enough to prevent surface oscillations. Mean square error gives poor insight as to what these regularization values should be; however, practice indicates that $\lambda = 0.1$, $\rho \geq 0.25$ work well.

7. FURTHER RESEARCH

The fitting methods described above treat all data points equally. This results in control points being assigned to relatively flat regions that do not need them. The fact that the optimum fit ratio is not the same as the ratio of the head scan data also suggests that a large amount of redundant information exists in the data sets. Research is under way to use adaptive methods to concentrate control points in areas of large surface fluctuations by removing unneeded rows and columns from the data set.

Other areas of research include fitting non-uniform and non-uniform/rational surfaces to the data sets. These should provide better preservation of discontinuities. Also the concept of mean square error is inadequate for describing surface errors related to surface oscillations. Research into describing the surface in frequency terms could greatly improve optimization decisions for regularization terms.

8. REFERENCES

1. J. Foley, A. vanDam, S. Feiner and J. Hughes, *Computer Graphics: Principles and Practice*, Addison-Wesley Publishing Co., Reading, Massachusetts, 1990.

2. G. Farin, *Curves and Surfaces for Computer Aided Geometric Design: A Practical Guide*, Academic Press, Inc., Boston, Massachusetts, 1993.

3. D.F. Rogers and J.A. Adams, *Mathematical Elements for Computer Graphics*, 2nd Edition, McGraw-Hill Publishing Co., New York, 1990.

4. P. Lancaster and K. Salkauskas, *Curve and Surface Fitting: An Introduction*, Academic Press, Boston, Massachusetts, 1990.

5. W.H. Press, S.A. Teukolsky, W.T. Vetterling and B.P. Flannery, *Numerical Recipes in C*, Cambridge University Press, New York, N.Y., 1992.

6. T. Poggio, V. Torre and C. Koch, "Computational vision and regularization theory", *Nature*, Vol. 317, pp. 314-319, 1985.

7. D. Terzopoulos, "Regularization of inverse visual problems involving discontinuities", *IEEE Trans. Patt. Anal. Machine Intell.*, Vol. 8, No. 6, p. 413-424, 1986.

8. S.S. Sinha and B.G. Schunck, "A Two-Stage Algorithm for Discontinuity-Preserving Surface Reconstruction", *IEEE Trans. Patt. Anal. Machine Intell.*, Vol. 14, No. 1, p. 36-55, 1992.

9. M. D. Corbin, "Surface Fitting of Large Data Sets", In preparation.

10. M. D. Corbin, MSEE Thesis, Ohio University, In preparation.

Edward Piepmeier report unavailable at time of publication.

MAINTAINING SKILLS AFTER TRAINING: THE ROLE OF OPPORTUNITY TO PERFORM TRAINED TASKS ON TRAINING EFFECTIVENESS

Miguel A. Quiñones
Assistant Professor
Department of Psychology

Rice University
6100 Main Street
Houston, TX 77005

Final Report for:
Summer Research Extension Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Rice University

December 1995

MAINTAINING SKILLS AFTER TRAINING: THE ROLE OF OPPORTUNITY TO PERFORM TRAINED TASKS ON TRAINING EFFECTIVENESS

Miguel A. Quiñones
Assistant Professor
Department of Psychology
Rice University

Abstract

Past research suggests that a critical aspect of training effectiveness is the extent to which trainees receive opportunities to perform trained tasks on the job. Individuals receiving such opportunities are more likely to retain the skills acquired during training and perform at higher levels than those not receiving many opportunities. A related, but relatively unexplored, consequence of differential opportunities to perform trained tasks is their effect on career outcomes such as assignments, promotions, and turnover. Early career experiences are likely to affect recent trainees' decision to remain in the organization and the rate at which they progress through the organizational hierarchy. The present study investigated the extent to which differential opportunities to perform as well as other pre- and post-assignment factors were related to reenlistment decisions at the end of an Airman's first term. Regression results suggest that attitudes formed during technical training as well as the level of support received during the first year on the job influenced Airmen's intentions to remain in the Air Force. Furthermore, intentions to remain in the Air Force predicted reenlistment decisions at the end of the first term (four years later). Implications for training effectiveness are discussed.

MAINTAINING SKILLS AFTER TRAINING: THE ROLE OF OPPORTUNITY TO PERFORM TRAINED TASKS ON TRAINING EFFECTIVENESS

Miguel A. Quiñones

Introduction

Recent technological and demographic trends suggest that employee training will become an increasingly critical component of an organization's competitive advantage (Goldstein & Gilliam, 1990; Rosow & Zager, 1988). Given these trends, it is important that the training programs employed by organizations have the desired result back in the workplace. Unfortunately, research suggests that only about 10% of training program expenditures result in performance changes on the job (Georgenson, 1982). Part of the problem stems from the lack of adequate program evaluation and the focus on training methods and techniques which employ the most recent "bells and whistles" while ignoring the factors which facilitate or impede the successful implementation of learned skills (Goldstein, 1993; Ford, Quiñones, Sego, & Sorra, 1992; Kraiger, Ford, & Salas, 1993).

Past research examining the "transfer problem", or lack of adequate implementation of trained skills to the job, have tended to focus on training design issues such as identical elements, task repetition, and training delivery (see Baldwin & Ford, 1988; Cormier & Hagman, 1987; Quiñones, Ford, Sego, & Smith, 1995). Ford, et al. (1992) suggested that a critical aspect of training effectiveness is the extent to which trainees receive sufficient experience or opportunities to perform trained tasks on the job. Their results found wide differences in the trained tasks performed by trainees months after leaving the training program. In addition, Ford, et al. (1992) as well as subsequent research (e.g., Ford, Smith, Sego, & Quiñones, 1993; Quiñones, et al., 1995) have identified a number of individual, work context, and organizational factors related to a trainee's level of experience or opportunity to perform trained tasks. Finally, Quiñones and Ford (1993) found a positive relationship between the number of times trainees performed trained tasks in the first eight months on the job, and their level of performance on that task. This relationship existed for over 60% of the tasks examined even after accounting for differences in general cognitive ability (AFQT) and initial level of learning.

The Opportunity to Perform

Ford et al. (1992) adopted a multidimensional perspective to define the opportunity to perform (see also Quiñones, Ford, & Teachout, in press). They defined opportunity to perform as the extent to which trainees are provided with or actively seek work experiences relevant to the tasks for which they were trained. Specifically, Ford et al. (1992) defined the "opportunity to perform" construct as consisting of at least three relatively independent dimensions: (1) breadth, or the number of trained tasks performed on the job; (2) activity level, or the number of times trained tasks are performed on the job; and (3) task type, or the difficulty and/or criticality of the trained tasks that are actually performed on the job.

Initial data collected on approximately 180 graduates from the AGE Technical Training School at Chanute AFB showed that four months after leaving technical training, trainees were only performing about half of the trained tasks sampled in the survey. The results also revealed wide differences in the type of tasks performed by the airmen as well as the number of times these tasks were being performed.

Factors Affecting the Opportunity to Perform

Initial analyses reported in Ford, et al. (1992) and Quiñones, et al. (1995) have identified a number of individual, work context, and organizational factors related to differences in opportunities to perform trained tasks. These include cognitive ability (AFQT), self-efficacy, career motivation, supervisor and peer support, supervisory attitudes, and organizational type (MAJCOM). From a training effectiveness viewpoint, these are all factors that can enhance or diminish the long term impact of training.

Turnover and Training Effectiveness

Analyses of the AGE data set have focused on opportunity to perform as an outcome variable. Preliminary evidence suggests that opportunity to perform can have an impact on various outcomes of interest such as performance (Quiñones & Ford, 1993) and career attitudes (Quiñones, et al., 1995). Both of these outcomes are important indicators of training effectiveness.

However, because the Air Force spends large sums of money training recruits, it is hoped that these recruits will remain in the military and utilize their skills in a productive manner. An effective training system requires that the gains due to training be realized in the field. Thus, factors which prevent individuals from applying their skills on the job and lead them to leave the military are diminishing the effectiveness of the entire training system.

The AT&T Management Progress Studies found that early career experiences were critical for the long term retention and success of managers (Howard & Bray, 1988). The career progression and mentoring literature has also identified early experiences, and the roles played by mentors (supervisors and peers), as key variables affecting the long term outcomes of individuals in organizations (cf. Kram, 1985; Noe, 1988). A related study highlighting the importance of training experiences on trainee outcomes was reported by Tannenbaum, Mathieu, Salas, and Cannon-Bowers (1991). These authors examined the extent to which a trainee's expectations regarding training were met and the negative consequences of unmet expectations (see also Quiñones, 1995). Opportunity to perform can be viewed as a measure of the extent to which trainees' expectations regarding their future careers are met. The extent to which trainees do not perform the tasks for which they were trained can result in negative career attitudes and turnover.

Other Factors Affecting Turnover

Past research on turnover has found that affective variables such as job satisfaction, commitment, and career motivation are related with intentions to quit and subsequent turnover (Hom & Hulin, 1981; Miller, Katerberg, & Hulin, 1979). Because training is often the first activity employees engage in when they enter an organization (Feldman, 1989), it can be the source of initial job attitudes that percolate down to subsequent turnover decisions. In addition, factors in the posttraining environment such as workgroup support can influence a trainee's experience with trained tasks as well as their attitudes toward the job (Ford et al., 1992; Quiñones et al., 1995).

Study Purpose

This research builds upon the literature on training effectiveness by examining the long term outcomes of experience or opportunity to perform as well as other pre- and post-assignment factors. As past research suggests, opportunities to perform trained tasks can enhance training effectiveness by maintaining and improving trainee's

skill and performance levels after training. However, there are other potential outcomes of differential opportunities to perform trained tasks that can affect training effectiveness. These include, among others, a trainee's career path and promotion schedule as well as their willingness to re-enlist after their initial term is completed.

The present study focused on reenlistment decisions as a critical dependent variable. Although other studies have examined the factors associated with positive reenlistment decisions, this study adds to that literature by examining the role that pre- and post-assignment factors and experiences play in reenlistment decisions. These factors were viewed as being important because they are potentially under the control of supervisors and training designers. Thus, any negative consequences associated with these factors can suggest interventions designed to retain a larger percentage of trainees.

Method

Sample. This study examined the pre- and post-training experiences of 255 Air Force recruits participating in the Aerospace Ground Equipment (AGE) Airman Basic-in-Residence (ABR) course in Chanute AFB from 3/89 to 10/89. The AGE ABR course consists of 18 weeks of instruction regarding powered and nonpowered equipment. Past research has found that the course instructs trainees on the performance of 99 tasks (Ford & Sego, 1990). However, the opportunity to perform measures focused on 34 of the 99 tasks (see Ford et al., 1992). All participants in this study were required to make reenlistment decisions by the time this data was collected.

The data available from previous studies contains information gathered at the end of training but prior to trainees receiving their first assignment. Key variables collected at this stage include training grades, AFQT scores, locus of control, career motivation. Trainees and their supervisors were then sent surveys at four, eight, and twelve months after training. These surveys contained the opportunity to perform measures as well as self-efficacy, peer and supervisor support, work flow, attitudes toward training, supervisor attitudes, and several variables describing the specific AGE shop (see Ford et al., 1992; Ford et al., 1993; Quiñones, et al., 1995). Additional data collected for this projects tracks the career progression of this original sample. The date contains variables such as re-enlistment status, current rank, awards, performance reviews, and additional specialty training received.

Analyses. The analyses performed for this project were aimed at examining the role of pre- and post-assignment factors on reenlistment decisions. Specifically, the extent to which attitudes developed during training, as well as those developed during the first year on the job, influence reenlistment decisions. Furthermore, the extent to which opportunities to perform trained tasks influenced attitudes and reenlistment decisions were also examined. Figure 1 presents the conceptual model driving the current research.

The specific variables examined in this study were divided into pre- and post-assignment clusters. The pre-assignment variables included training grades, AFQT (see United States Department of Defense, 1984), locus of control, and training reactions. These variables were hypothesized to affect a trainee's level of career motivation prior to getting their first assignment. Post-assignment variables examined included opportunity to perform (breadth, activity level, and task type), supervisory and peer support, as well as pre-assignment career motivation. In addition, final training grades and AFQT were used as covariates in order to determine the unique contributions of the other factors. These variables were hypothesized to influence a trainee's intentions to remain in the Air Force. Intentions to remain were gathered at the end of the first year after training. Finally, intentions to remain at the end of the first year were hypothesized to influence reenlistment decisions four years later (at the end of the five-year commitment). In addition, pre-assignment career motivation, supervisor and peer support, as well as task type were used as predictors in the equation to covary out their effects from those of intentions.

The statistical methods employed for these analyses included hierarchical multiple regression (e.g., Cohen & Cohen, 1983) and logistic regression (e.g., Hosmer & Lemeshow, 1989; Liao, 1994). The procedures allow for the examination of the independent contribution of several variables in predicting a dependent variable of interest. The main difference rests upon the nature of the dependent variable. Multiple regression was used for continuous (interval or ratio) dependent variables whereas logistic regression was used for predicting a dichotomous dependent variable such as reenlistment status.

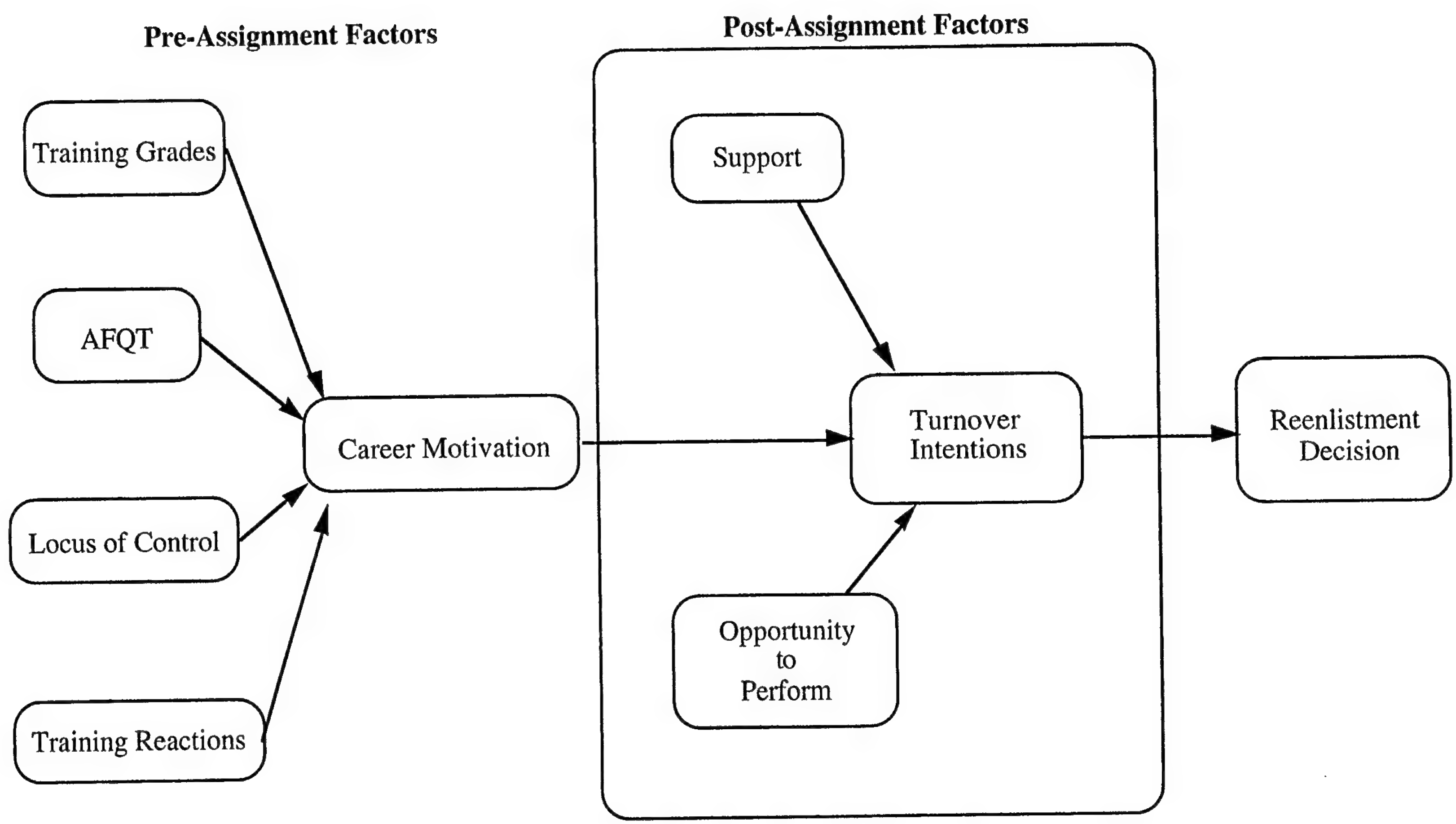


Figure 1: Conceptual Model

Results

Descriptive Data. Table 1 presents the means, standard deviations, and intercorrelations for the variables examined. Perhaps of most interest is the reenlistment variable indicating that 53.7% of the original sample decided to leave the Air Force at the five-year reenlistment period. The remainder of the variables appeared to have adequate distributional properties necessary for regression analyses.

Pre-Assignment Factors. A regression model was estimated predicting career motivation at the end of training but prior to being assigned to a duty station. The independent variables examined were training grades, AFQT, locus of control, and training reactions (see Figure 1). Table 2 presents the results of this analysis. The results show that training reactions were positively related to career motivation at the end of training ($\beta = .52$, $p < .01$). Trainees who liked the training program reported being more excited about their career in the AGE field. AFQT scores were found to be negatively related to career attitudes ($\beta = -.15$, $p < .01$). Individuals scoring high on the AFQT composite of the ASVAB reported lower levels of career motivation relative to their low scoring counterparts. All variables accounted for 32% of the variance in career motivation scores.

Post-Assignment Factors. The second set of analyses were aimed at identifying workplace factors which contributed to trainees' intentions to remain in the Air Force. Supervisor and peer support as well as the three opportunity to perform factors (breadth, activity level, and task type) were used as independent variables in this analysis. In addition, AFQT and training grades were entered as covariates. Entering these two variables also determines the extent to which the higher scoring trainees tend to have higher intentions to remain.

Table 3 presents the results of this analysis. Career motivation as well as peer and supervisory support were found to be positively related to intentions to remain in the Air Force ($\beta = .17$, $p < .01$ and $\beta = .25$, $p < .01$, respectively). None of the other variables in the model were significantly related to intentions to remain in the Air Force. All variables accounted for 12% of the variance in intentions to remain in the Air Force.

Table 1

Means, Standard Deviations, and Intercorrelations of Study Variable.

Variable	Mean	SD	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
1. Training Grades	90.34	5.03	1.00										
2. AFQT	60.77	15.63	.32**	1.00									
3. Locus of Control	4.52	1.12	.21**	.20	1.00								
4. Training Reactions	40.02	5.39	-.03	.09	.30**	1.00							
5. Career Motivation	57.53	7.65	.03	-.02	.21**	.57**	1.00						
6. Support	26.88	5.31	-.04	-.03	.08	.15**	.16**	1.00					
7. Breadth	19.77	5.86	.11*	.00	-.05	-.11	.02	.15**	1.00				
8. Activity Level	311.41	284.02	.13**	.00	-.03	-.07	.01	.05	.35**	1.00			
9. Task Type	15.50	3.50	.06	.00	.17**	.08	.21**	.54**	.37**	.18**	1.00		
10. Intentions to Remain	8.36	3.52	-.03	-.06	.12*	.13*	.19**	.25**	.07	.07	.25**	1.00	
11. Reenlistment Status	.46	.50	.00	.01	.10	.02	.07	.20**	.14**	.09	.14**	.33**	1.00

* $p < .05$

** $p < .01$

Table 2

Regression Analyses Predicting Career Motivation from Pre-Assignment Factors.

VARIABLE	B	SE B	β	t
Training Grades	.04	.08	.03	.49
AFQT	-.06	.03	-.15	-2.69**
Locus of Control	.61	.35	.10	1.73
Trainee Reactions	.72	.07	.52	9.64**

$R^2 = .32$

$F(4,258) = 30.13$, $p < .001$

** $p < .01$

Table 3

Regression Analyses Predicting Intentions to Remain in the Air Force from Post-Assignment Factors.

VARIABLE	B	SE B	β	t
Training Grades	.00	.05	.00	-.04
AFQT	-.01	.01	-.04	-.61
Career Motivation	.08	.03	.17	2.69**
Support	.09	.03	.25	2.99**
Activity Level	.00	.00	.08	1.20
Breadth	.00	.04	-.01	-.18
Task Type	.02	.05	.03	.34

$R^2 = .12$

$F(4, 247) = 4.94$, $p < .001$

** $p < .01$

Reenlistment Decisions. Finally, a logistic regression analysis was conducted in order to predict which individuals were more likely to reenlist at the 5-year mark. The independent variables examined included intentions to remain in the Air Force, supervisor and peer support, career motivation, and task type. Previous research suggests that task type may affect this decision so it was decided to include this variable in the model (see Quiñones, et al, 1995).

The results of the logistic regression revealed that the equation significantly predicted those individuals who reenlisted ($\chi^2 = 325.46$, $p < .01$). Specifically, the equation successfully predicted 68.2% of the cases. The results also showed that higher levels of supervisor and peer support as well as intentions to remain in the Air Force increased the probability that an individual reenlisted at the 5-year period ($b = .07$, $p < .05$ and $b = .20$, $p < .01$, respectively).

Conclusions

The purpose of this study was to build upon past research on transfer of training by examining the extent to which differential opportunities to perform trained tasks as well as other pre- and post-assignment factors predict reenlistment decisions among a group of graduates from the AGE ABR course. The results of several regression analyses suggest that reenlistment decisions can be reasonably predicted (68.2% accuracy) by a number of factors. Specifically, the results suggest that individuals who reacted positively to training and had lower AFQT scores left Chanute AFB with higher career motivation when compared to their counterparts. It was also found that this level of career motivation as well as supervisory and peer support in the first assignment was related to trainees' intentions to remain in the Air Force during their first year on the job. Perhaps the most interesting finding was the fact that intentions to remain in the Air Force as well as level of support during the first year predicted reenlistment decisions four years later.

Several conclusions and implications can be drawn from these findings. First, it is clear that recruits form their opinions about their jobs and the Air Force early in their careers. The initial training program is clearly a source of information which they use to make judgments about their careers. It is interesting to note that high

ability (AFQT) individuals left training with lower levels of career motivation. Perhaps being exposed to the reality of the AGE career field left these individuals with the impression that they would not be intellectually challenged. This suggests that there should be a top cutoff of AFQT scores developed for assignment into the AGE career field.

In addition to AFQT scores, trainee reactions were important in developing high levels of career motivation. Clearly trainees are not passive recipients of information during training. Trainees form opinions about the quality of the training program which translate into opinions about their career field. Clearly training designers must consider trainee reactions when modifying training programs.

The level of career motivation developed as a result of the training experience was shown to carry over into the field. Along with the level of support received, career motivation determined the extent to which trainees wished to remain in the Air Force. Clearly, these early intentions remain with the trainee throughout their final years on the job and influence their reenlistment decisions. These findings suggest that early interventions are critical for retaining Air Force personnel. Interventions designed to identify and ameliorate negative attitudes towards an Airman's career can go a long way toward reducing the level of turnover found in this study. If the Air Force continues to spend large sums of money training individuals on a variety of technical skills, these individuals' attitudes must be considered if the Air Force is to capitalize on its investment.

REFERENCES

Baldwin, T.T. & Ford, J.K. (1988). Transfer of training: A review and directions for future research. Personnel Psychology, 41, 63-105.

Cormier, S.M. & Hagman, J.D. (1987). Transfer of Learning. San Diego: Academic Press.

Cohen, J. & Cohen, P. (1983). Applied multiple regression/correlation analysis for the behavioral sciences (3rd ed). Hillsdale, NJ: Lawrence Erlbaum.

Feldman, D.C. (1989). Socialization, resocialization, and training: Reframing the research agenda. In I.L Goldstein and Associates (Eds.), Training and Development in Organizations. San Francisco, CA: Jossey-Bass.

Ford, J.K. & Sego, D.J. (1990). Linking training evaluation to training needs assessment: A conceptual model (AFHRL TP 90-69). Brooks AFB, TX: Air Force Human Resources Laboratory.

Ford, J.K., Quiñones, M.A., Sego, D.J., & Sorra, J.S. (1992). Factors affecting the opportunity to perform trained tasks on the job. Personnel Psychology, 45, 511-527.

Ford, J.K., Smith, E.M., Sego, D.J., & Quiñones, M.A. (1993). Impact of Task Experience and Individual Factors on Training Emphasis Ratings. Journal of Applied Psychology, 78, 583-590.

Georgenson, D.L. (1982). The problem of transfer calls for partnership. Training and Development Journal, 36, 75-78.

Goldstein, I.L. (1993). Training in organizations: Needs assessment, development, and evaluation (3rd ed). Monterey, CA.: Brooks/Cole.

Goldstein, I.L. & Gilliam, P. (1990). Training systems in the year 2000. American Psychologist, 45, 134-143.

Hosmer, D.W. & Lemeshow, S. (1989). Applied Logistic Regression. New York : Wiley.

Howard, A. & Bray, D.W. (1988). Managerial lives in transition: Advancing age and changing times. New York: The Guilford Press.

Hom, P.W. & Hulin, C.L. (1981). A competitive test of the prediction of reenlistment by several models. Journal of Applied Psychology, 66, 23-39.

Kraiger, K., Ford, J.K., & Salas, E. (1993). Application of cognitive, skill based, and affective theories of learning outcomes to new methods of training evaluation. Journal of Applied Psychology, 78, 311-328.

Kram, K.E. (1985). Mentoring at work: Developmental relationships in organizational life. Glenview, Ill.: Scott, Foresman.

Liao, T.F. (1994). Interpreting probability models. Thousand Oaks, CA: Sage.

Miller, H.E., Katerberg, R., & Hulin, C.L. (1979). Evaluation of the Mobley, Horner, and Hollingsworth model of employee turnover. Journal of Applied Psychology, 64, 509-517.

Noe, R.A. (1988). An investigation of the determinants of successfully assigned mentoring relationships. Personnel Psychology, 41, 457-480.

Quiñones, M.A. (1995). Pretraining context effects: Training assignment as feedback. Journal of Applied Psychology, 80, 226-238.

Quiñones, M.A., Ford, J.K., Sego, D.J., & Smith, E.M. (1995). The effects of individual and transfer environment characteristics on the opportunity to perform trained tasks. Training Research Journal, 1, 29-48.

Quiñones, M.A., Ford, J.K., & Teachout, M. (in press). The relationship between work experience and job performance: A conceptual and meta-analytic review. Personnel Psychology

Quiñones, M.A. & Ford, J.K. (1993, May). Individual and task experience effects on transfer of training: A field study. Paper presented at the 8th Annual Conference of the Society for Industrial and Organizational Psychology, San Francisco, California.

Rosow, J.M. & Zager, R. (1988). Training - The competitive edge. San Francisco: Jossey-Bass.

Tannenbaum, S.I., Mathieu, J.E., Salas, E., & Cannon-Bowers, J.A. (1991). Meeting trainee's expectations: The influence of training fulfillment on the development of commitment, self-efficacy, and motivation. Journal of Applied Psychology, 76, 759-769.

United States Department of Defense (1984). Armed Services Vocational Aptitude Battery (ASVAB): Test Manual (DoD Publication No 1304.12AA). Washington, DC: United States Military Processing Command.

Gary Riccio report unavailable at time of publication.

Documentation of Separating and Separated Boundary Layer Flow, For Application

Wayne Shebilske
Professor
Department of Psychology

Texas A&M University
College Station, TX 77845

Final Report for:
Summer Research Extension Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base
Washington, D.C.

and

Armstrong Laboratory

December 1995

DOCUMENTATION OF SEPARATING AND SEPARATED BOUNDARY LAYER FLOW, FOR APPLICATION

Wayne Shebilske
Professor
Department of Psychology
Texas A&M University

Abstract

The chapter presents a componential iterative learning phase model, which guides the three experiments presented in this final report. The proposal for global analyses of Total Scores in the present research was guided by a traditional learning phase model, which applies phases globally to a whole task. The model assumes that tasks go through one cycle starting with a Declarative Knowledge phase, which is dominated by explicit processes, transitioning through Knowledge Compilation, and ending in a Procedural Knowledge phase, which is dominated by implicit processes. The present research helped us understand that the traditional model does not apply to complex sensorimotor tasks, which have important explicit and implicit processes throughout skill development. By manipulating variables that are known to have different effects on explicit and implicit processes, the present experiments show the importance of componential analyses, which reveal separate effects on components that are dominated by either implicit or explicit processes. Contrary to the traditional model , the present experiments emphasized the importance of implicit processes early in training. These experiments will be published along with other experiments that have been run in the TRAIN laboratory showing the increasing importance of explicit processes late in training.

FINAL REPORT : Executive Summary (Subcontract No. 95-0814; RF48878100009)

The support of the present grant is acknowledged in the Appended chapter in an internationally respected book, Shebilske, W. L., Goettl, B., & Regian, J. W. (in press). Individual and group protocols for training complex skills in laboratory and applied settings. In Gopher, D. & Koriat, A. (Eds.) Attention and Performance XVII: Cognitive regulation of performance: Interaction of theory and application. Hillsdale, NJ: Earlbaum. The chapter presents a componential iterative learning phase model, which guides the three experiments presented in this final report. The proposal for global analyses of Total Scores in the present research was guided by a traditional learning phase model, which applies phases globally to a whole task. The model assumes that tasks go through one cycle starting with a Declarative Knowledge phase, which is dominated by explicit processes, transitioning through Knowledge Compilation, and ending in a Procedural Knowledge phase, which is dominated by implicit processes. The present research helped us understand that the traditional model does not apply to complex sensorimotor tasks, which have important explicit and implicit processes throughout skill development. By manipulating variables that are known to have different effects on explicit and implicit processes, the present experiments show the importance of componential analyses, which reveal separate effects on components that are dominated by either implicit or explicit processes. Contrary to the traditional model , the present experiments emphasized the importance of implicit processes early in training. These experiments will be published along with other experiments that have been run in the TRAIN laboratory showing the increasing importance of explicit processes late in training.

FINAL REPORT (Subcontract No. 95-0814; RF48878100009)

Introduction

The present analysis of inter-lesson intervals and inter-lesson processing activity emphasizes the changing implicit and explicit cognitive demands as skill develops in highly interactive components of hierarchically structured tasks. Hierarchically structured tasks have basic component skills that must be learned before their higher level, complex component skills can be learned. As skill develops, performance of these basic components must be integrated with the more complex components to form integrated, strategic responses. This complex integration of responses, or "response ensembles" (Gopher 1987), is what is meant by high componential interactivity. An example is coordinating the many motor, perceptual, and cognitive skills required to fly an airplane. Response ensembles include implicit and explicit processes. Implicit learning is not verbalizable. It is measured by an increase in performance on a criterion task associated with an inability by the learner to relate how they improved. Examples include classical cconditioning and learning artificial grammars (Reber, 1993) . In contrast, explicit learning is verbalizable and open to conscious hypothesis testing or pattern recognition (Seger, 1994). The present research investigates the effects on implicit and implicit processes of practice lesson spacing and explicit inter-lesson processing activity. Explicit cognitive processes include elaborative rehearsal , problem-solving, and mental rehearsal . A lesson consists of approximately 30 minutes of practice and test games on a representative analogue of criterion tasks with high componential interactivity.

For tasks with low componential interactivity, past research indicates a stable or decreasing cognitive demand. These findings were explained by a theory of skill learning in which cognitive abilities predict performance early in learning and psychomotor ability predicts perfomance late in learning (Ackerman, 1987). In contrast, Corrington and Shebilske (1995) proposed that explicit cognitive demands may increase as training proceeds for tasks with high componential interactivity and a hierarchical structure. In addition, they proposed that implicit, automatic processing may occur early in skill development for some subcomponents of complex tasks. These proposals were further developed and formalized as the componential iterative model of skill learning (Shebilske, Goettl, & Regian, 1996). This model has guided past research on subjective reports, attention-ability relationships, and intelligence-performance relationships. The model's emphasis on the interplay between implicit and explicit processes thoughout learning has also guided the present research. The emphasis was especially important in light of previous literature suggesting that highly massed practice, with no inter-lesson interval, and inter-lesson explicit processing may have opposite effects on implicit and explicit learning. Specifically, massed practice lessons facilitate implicit learning and hinder explicit learning, while explicit inter-lesson processing facilitates explicit learning and may hinder implicit learning.

In the present research, three laboratory experiments manipulated spacing of practice lessons, with either 0 or 25 minute intervals between lessons, and opportunites for inter-lesson processing, which were various kinds of explicit processing or processing of a filler task that preempted explicit processing. The dependent variables were performance on subscores and Total Scores for test games. The purpose was to analyze effects of both

massed practice and explicit inter-lesson processing on implicit and explicit learning. These experiments, employing a video game-like research tool, Space Fortress, extend previous research on inter-lesson spacing of practice on a complex, sensorimotor task.

Experiment 1

The goal of the first experiment was to analyze effects of inter-lesson interval (0 vs. 12 min.) and explicit inter-lesson processing (explicit vs. filler) on implicit and explicit learning processes, which were reflected by sub-component scores. A pilot study suggested that Control and Velocity subscores reflect primarily implicit and explicit processes and that Speed and Points subscores are primarily indicators of explicit strategic processes. None of the sub-components exclusively indicate either implicit or explicit processes. Some explicit influences on the Control score involve motoric behaviors, such as letting the joy stick center itself, and some involve strategic behaviors, such as letting mines come to the ship. Similarly, although Speed and Points scores seem to be highly influenced by explicit strategic behaviors, they are also influenced by implicit processes. Implicit influences include all implicit motor skills that are necessary for executing verbalizable strategies as well as non-verbalizable strategic habits.

An empirical question addressed in the present study is whether dominant influences are strong enough to tease apart separate effects of variables on implicit and explicit processes. If they are, Control and Velocity should be improved by inter-lesson massing, while Speed and Points should be improved by inter-lesson spacing and by explicit inter-lesson processing. The predictions for Control and Velocity are made with more confidence based on the assumption that higher skills in a hierarchy influence lower skills more than lower skills influence higher ones. Accordingly, Control and Velocity,

which are dominated by implicit skills near the top of the hierarchy, are purer indicators of implicit processes. In contrast, Speed and Points, which include many explicit skills at lower hierarchical levels, are less pure indicators of explicit processes.

Method

Participants

The sample consisted of 102 students from Texas A&M university who received credit in an Introductory Psychology course for their participation. Selection was based on an aiming screening task (Mane & Donchin, 1989) and a survey of video game experience. Subjects were not permitted to participate if they failed to obtain a minimum aiming score of 780 points or if they reported currently playing video games more than 20 hours per week.

Apparatus and Tasks

A laboratory room was equipped with four tables, each with an IBM PC/AT compatible computer, color monitor, joystick (Flightstick by CH Products), a three-button mouse, and two right-handed chair-desks. Two straps secured the joystick to the chair-desk and the subject's arm was supported comfortably in a horizontal position perpendicular to the joystick. The mouse was placed on the other chair-desk. The Space Fortress task used in the present study was a version of the game developed by Gopher et al. (1994) .

Space Fortress

In Space Fortress, the subject sits facing a computer screen. Displayed in the middle of the screen is a fortress. Next to it is a ship, which the subject controls via a

joystick. The goal of the game is to fire ship missiles at the fortress in an attempt to destroy it.

The ship flies in a no friction environment. Forward movements of the joystick cause the ship to accelerate. Lateral movements cause the ship to rotate in the direction that the joystick is moved. In this environment, the ship will continue to fly in the direction in which it is pointing unless it is rotated and thrust is applied. The ship will also fly off the screen. When this occurs, it is in "hyper space" and will reappear on the opposite side of the screen.

Destruction of the fortress is achieved by making it vulnerable. This is done by hitting it 10 times with missiles. These hits must be at least 250 msec. apart. After the fortress is made vulnerable, it can be destroyed with a double shot--two shots fired within 250 msec. of each other. While attempting to destroy the fortress, the subject must also defend against the fortress shells and foe mines. Mines appear every 4 seconds on the screen. The mines are either "friend" or "foe". To determine whether the mine is friend or foe, the subject must see a letter at the bottom of the screen. If the letter matches one of three letters presented prior to the game, then it is a friend. Otherwise it is a foe. If the mine is a foe, then the subject must switch weapons systems by pressing a mouse button twice within a 250-400 msec. window. The foe mine can then be destroyed. If the mine is friendly, then no switching of weapons systems is required. The subject simply shoots the mine, which then becomes "energized" and capable of scoring a hit against the fortress. While mines are on the screen the ship is ineffective in making a hit on the fortress.

The ship is damaged when a mine or a fortress shell strikes it. When the ship is damaged four times the ship is destroyed and the game is reset back to its starting

configuration. Points are continually added or subtracted during the course of the game. For example, subjects must track their weapons supply and bonus opportunities. They start with 100 missiles. When their supply runs out, they are able to shoot, but 3 points are subtracted for each shot. During the course of the game, more missiles or a bonus score can be obtained when a dollar sign appears on the screen for the second time. (Other symbols are shown between the presentations of the successive dollar signs.) Subjects then press one button to obtain more missiles or another button to receive bonus points.

A panel at the bottom of the screen displays four subscores. A velocity score increases when the ship flies below a threshold velocity. A control score increases when the ship is flown within the hexagon surrounding the fortress; it decreases when the ship leaves the screen and goes into hyperspace. A speed score increases when mines are identified and dealt with properly. A points score increases when the mines and fortress are destroyed; it decreases when the ship is damaged or destroyed. A total score, which is a sum of the other four scores, is displayed at the end of the game.

Aiming Screening Task

In the Aiming Screening Task (Mane and Donchin, 1989), a stationary ship lies in the middle of the computer screen. The ship can be rotated left by moving the joystick to the left and rotated right by moving the joystick to the right. Mines, in the shape of blue diamonds, appear on the screen. The object of the task is to shoot as many of the mines as possible during the course of a game. Each game lasted 1-minute and subjects performed 3 games. If a subject is unable to shoot a mine it disappears after 4-seconds. After a

mine disappears, either through its destruction or because a subject is unable to destroy it, another mine appears after a 1-second interval.

Inter-lesson Processing Exercise

The inter-lesson processing exercise was an open-ended essay question. It asked participants to "use this time between lessons to write down your thoughts pertaining to the previous lesson of Space Fortress or to the upcoming lesson". Its intent was to record any spontaneous thoughts concerning the task without biasing them in any way.

Probe of Top-Down Processing

The Probe of Top-Down Processing asked subjects to "describe your efforts during the last game to improve your proficiency". This probe occurred after 1 practice game following Lesson 4. Participants believed they were to complete an entire 5th Lesson. All conditions were given the probe. The intent of the instrument was to capture any differences between groups in top-down processing related to Space Fortress.

Anagrams

The massed and spaced-anagram conditions performed 3 anagram filler tasks. Each anagram filler task consisted of 8 words. Subjects were instructed to produce as many words as possible, each composed of a subset of the letters of the 8 words. Anagrams for each word was not a requirement. Subjects were instructed to move to the next word if they were unable to produce any words. Pluralized words, slang words, proper nouns, and obsolete or archaic words were not allowed. Only one form of each word was allowed; for example, only bottle or bottles, but not both.

The 8 words in anagram filler task A were concertina, premises, desperado, musically, underworld, telephone, dungarees, and toaster. Words in anagram filler task B

were scramble, plastic, publisher, composer, wisecrack, obsolete, astronaut, an alphabet. Words in anagram filler task C were spectacle, methodic, soldiers, capitalist, seraglio, blowouts, gold-diggers, and chameleon. Subjects were equally and randomly assigned to filler task presentation order. The intent of the tasks was to occupy processing resources between Space Fortress lessons.

Confidence and Alertness Questionnaire

Composed of 12 items, answers give retrospective subjective reports on confidence, alertness, and sleepiness. Confidence in performance relative to other trainees, confidence in knowledge of rules and strategies,and confidence in ability to improve on the upcoming lesson are examples of confidence items. Trainees were asked how sleepy and how alert they were to evaluate these subjective states.

Experimental Design

Participants, assigned to one of three conditions, were matched on baseline Total Scores. A criterion of plus or minus 400 points was used to match subjects. Condition served as a between-subjects independent variable. The conditions were: (1) massed (0 minute interlesson interval) (2) spaced-filler (12 minute interlesson interval, during which a filler task was performed) and, (3) spaced-explicit (12 minute interlesson interval, during which an interlesson processing exercise was performed). This between-subjects independent was factorially combined with 2 within-subject independent variables: (a) Acquisition, Lessons 1-4 and (b) retention, Lesson 5. Time of day for training was counterbalanced across participants, who were equally and randomly assigned to morning, afternoon, or evening. The dependent variables were (a) sub-scores averaged over 2 test games per lesson, (b) Total Score averaged over 2 test games per lesson, and (c) answers

to questions in the Confidence and Alertness Questionnaire (CAQ) and the Probe of Top-Down Processing.

Procedure

Subjects signed an informed-consent form, filled out a survey of video-game experience, and then performed the aiming screening task. Subjects who passed these screening measures were randomly assigned to the three experimental conditions. Next, subjects watched a 17-minute instruction video before they played 4, 3-minute baseline games of Space Fortress. Following the baseline games, subjects had a 10 minute break before they watched a 5-minute "summary of instructions" video. After the summary of instructions, all subjects performed 4 lessons of Space Fortress, each consisting of 5, 3-minute practice games and 2, 3-minute test games. (The inter-lesson processing exercise and filler task were performed between Space Fortress Lessons 1 and 2, Lessons 2 and 3, and Lessons 3 and 4 by the 12-minute interval groups. The 0-minute interval group performed the filler task after the retention games). Following Lesson 4, subjects thought they were to perform Lesson 5. However, after 1 practice game, they were stopped and they were given 5 minutes to fill out the Probe Questionnaire. After the Probe, participants filled out the CAQ. After the CAQ, participants performed a retention lesson, which consisted of 2, 3-minute test games. To conclude, subjects were debriefed.

Results

Mixed, repeated measures analyses of variance (ANOVA) were conducted, with condition a between- subjects factor and acquisition and retention lessons within-subjects factors, on all test game sub-scores for Lesson 1 through Lesson 4 . Overall analyses

revealed main effects for Lesson on Control Score, F(4, 396) = 73.37, $p < .01$, Velocity Score, F(4,396) = 10.57, $p < .01$, Speed Score, F(4,396) = 55.49, $p < .01$, and Points Score, F(4, 396) = 126.11, $p < .01$. No significant interactions were indicated. A mixed repeated measures ANOVA was also conducted for test game Total Scores for Lesson 1 through Lesson 4. A main effect was found for Total Score, F(4, 396) = 124.31, $p < .01$. No significant interactions for Total Score were indicated. Planned comparisons revealed no significant differences between groups on Lesson 4 or retention test games, for either sub-scores or Total Scores. Extremely poor baseline and Lesson 4 test scores were evidenced in the study. This is atttibuted to low scores female participants (see Table 1). No significant differences between groups were evidenced for any items on the Confidence and Alertness Questionnaire. The matched group design allowed a second, totally within repeated measures analysis of variance (ANOVA) to be conducted, using the massed participant as the first observation, the spaced-filler participant as the second observation, and the spaced-explicit participant as the third observation for each dependent variable. This analysis revealed the same pattern of null results.

For the majority of the subjects, the explicit processing manipulation did not take. Subjective analyses by proctors revealed that most participants in the spaced-explicit condition did not utilize the 12 minutes given them. Often, time-on-task was approximately 2 to 3 minutes. Low word counts were consistent with these reports. As a result, most responses to the open-ended essays were too sparse to be of value. Fortunately, enough thoughtful answers were written to guide design of the inter-lesson processing exercise used in Experiment 2. Answers were also sparse on the Probe of Top-

Down Processing, indicating that the manipulation did not take for the majority of subjects.

Discussion

No effects of explicit inter-lesson processing or massed practice on sub-component scores were evidenced. Several problems with the design prevent definitive conclusions regarding these null results. These problems were: (1) Female trainees were used in this experiment. Past research indicates a large gender effect in Space Fortress, such that males perform significantly better than females. This served to reduce the power of the study. (2) Five practice games were performed instead of the traditional 8 practice games. This may have reduced the rate of learning, eliminating any potential advantages of massed practice or explicit inter-lesson processing for task components. (3) Open-ended questions failed to produce explicit inter-lesson processing activity. As discussed above, the manipulation failed to take. Guided questions may be required to produce explicit processing. (4) The anagram filler task may not have preempted cognitive processing related to Space Fortress. A more cognitively absorbing task may be required to prevent explicit inter-lesson processing related to the primary task. These problems were addressed in Experiment 2.

Experiment 2

The present experiment made changes to remedy 4 problems with Experiment 1. First, only male participants were used to reduce within-cell variance. Second, the number of practice games was increased from 5 to 8 to increase treatment effects. Third, guided questions were used to induce explicit interlesson processing. The interlesson processing exercises requested comparisons between worst and best games with respect to each game

component as well as problem diagnosis and solution planning with respect to components. Thoughtful responses to the open-ended essay in Experiment 1 were used as a basis for these questions. Fourth, a math solve and a keypress filler task were piloted. In this study, massed condition trainees performed the task at the end of the training session. Their results were used to evaluate the tasks as potential fillers to preempt explicit inter-lesson processing in subsequent experiments.

In Experiment 2, participants performed 3 Acquisition Lessons and 1 Retention Lesson on Space Fortress. Manipulated variables were spacing of practice, which were either 0 or 25 minute interlesson intervals, and opportunities for explicit processing, which were guided elaboration, problem-diagnosis, problem-solution, and mental rehearsal exercises between lessons or no explicit processing between lessons. The dependent variables were performance on sub-scores and Total Scores for test games.

The predictions are the same as those in Experiment 1. Inter-lesson massing and a lack of inter-lesson processing should improve Control and Velocity Scores, while inter-lesson spacing and inter-lesson explicit processing should improve Speed and Points Scores. The predictions for Control and Velocity are made with more confidence because they should be purer indicators of implicit processes, as discussed in Experiment 1.

Method

Participants

The sample consisted of 48 male students from Texas A&M university who received credit in an Introductory Psychology course for their participation. Selection was based on an aiming screening task (Mane & Donchin, 1989) and a survey of video game experience. Subjects were not permitted to participate if they failed to obtain a minimum

aiming score of 780 points or if they reported currently playing video games more than 20 hours per week.

Apparatus and Tasks

A laboratory room was equipped with four tables, each with an IBM PC/AT compatible computer, color monitor, joystick (Flightstick by CH Products), a three-button mouse, and two right-handed chair-desks. Two straps secured the joystick to the chair-desk and the subject's arm was supported comfortably in a horizontal position perpendicular to the joystick. The mouse was placed on the other chair-desk. The Space Fortress task used in the present study was a version of the game developed by Gopher et al. (1994). The Space Fortress task and the Aiming Screening Task (Mane & Donchin, 1989) are described in Experiment 1.

Inter-lesson Processing Exercises

The inter-lesson processing exercises were divided into 3 components--an explicit processing exercise, a self-diagnosis and planned solutions exercise, and a mental practice exercise. The explicit processing exercise asked participants to "compare your best and worst games, or parts of games, with respect to the following: (1) joystick manipulation for controlling ship velocity and position, (2) identifying and shooting mines, (3) shooting the fortress or being shot by the fortress, (4) collecting bonuses, (5) implementing optimal strategies. Each of these subsections was followed by space to write entitled Best: and space to write entitled Worst:. The self-diagnosis and planned solutions exercise contained the same sections as the elaboration exercise, only the initial instructions read: "describe your problems and possible solutions with respect to the following (consider changes in how you practice). Furthermore, each section was followed by space to write

called Problem(s): and a space to write called Solution(s):. Two minutes were allotted for each of these questions. Finally, instructions were read to the participants that asked them to visualize performing the solutions they devised. They visualized their game-playing for 3 minutes. To allow time for these exercises, the interlesson interval was increased in the present experiment, from 12 minutes to 25 minutes.

Progress Report

The Progress Report asked participants 4 questions: (1) State your initial ability and progress made in learning with respect to each of the following: (a) playing the first 30 seconds of a game, (b) using the display panels at the bottom of the screen, (c) controlling the ship's velocity and position, (d) responding to mines, (e) shooting the fortress and avoiding being shot by the fortress, (f) collecting bonus points. (2) The video-taped instructions described 5 optimal strategies for playing Space Fortress. Describe each strategy, state your initial ability to employ each one, and describe any progress that you made in learning to use each one. (3) Did you play practice games and test games the same way? Yes or No. If No, please explain. (4) Did you develop any special methods for learning the game during practice or any special strategies for playing the game during tests? Yes or No. If Yes, please explain. This report was used to indicate any differences between groups in the questions asked, to assess whether the processing exercises affected the participants approach to the task.

Learning Styles Questionnaire

The Learning Styles Questionnaire was administered to subjects to measure the degree to which a subject uses elaborative rehearsal when learning. It was used to assess

any interactive effects between learning style and practice spacing. The questionnaire used a 6-point Likert scale, with 1 labeled "strongly disagree" and 6 labeled "strongly agree".

Math Solve and Keypress Tasks

The Math Solve Task had subjects add and subtract numbers. These problems were presented on the computer monitor. Reaction times and errors were recorded. The Keypress Task displayed a letter on the computer screen. Again, reaction times and errors were recorded. These two tasks were alternated, one after the other, such that each task was performed 4 times during one inter-lesson interval. Thus, a session is 8 games long. Each task is 3 minutes long, so one lesson lasts 24 minutes. These tasks were performed by the massed group after the Progress Report. They were used for two purposes: (1) to make total time training equilavent for both groups and (2) to test them as potential filler tasks in future studies.

Experimental Design

Participants, assigned to one of two conditions, were matched on baseline Total Scores, using the same criterion as in Experiment 1. Condition served as a between-subjects independent variable. The conditions were: (1) massed (0 minute interlesson interval) and (2) spaced-explicit (25 minute interlesson interval, during which an inter-lesson processing exercise was performed). This between-subjects independent variable was factorially combined with 2 within-subjects independent variables: (a) Acquisition, Lessons 1-3 and (b) retention, Lesson 4. Time of day for training was counterbalanced across participants, who were equally and randomly assigned to morning, afternoon, or evening. The dependent variables were (a) sub-scores averaged over 2 test games per

lesson, (b) Total Score averaged over 2 test games per lesson, and (c) answers to questions in the Progress Report and Learning Styles Questionnaire.

Procedure

Subjects signed an informed-consent form, filled out a survey of video-game experience, and then performed the aiming screening task. Subjects who passed these screening measures were randomly assigned to the three experimental conditions. Next, subjects watched a 17-minute instruction video before they played 4, 3-minute baseline games of Space Fortress. Following the baseline games, subjects had a 3 minute break before they watched a 5-minute "summary of instructions" video. After the summary of instructions, all subjects performed 3 lessons of Space Fortress, each consisting of 8, 3-minute practice games and 2, 3-minute test games. (The inter-lesson processing exercises were performed between Space Fortress Lessons 1 and 2 and between Lessons 2 and 3 by the 25-minute interval groups. The 0-minute interval group performed the filler tasks after completing the Progress Report.) Next, all subjects filled out the Learning Styles Questionnaire. The interval between lesson 3 and the first retention lesson, during which subjects filled out the Learning Styles Questionnaire, was 3 minutes and was closely monitored. After the questionnaire, participants performed a retention lesson, which consisted of 2, 3-minute test games. Then subjects completed the Progress Report. To conclude, subjects were debriefed.

Results

Mixed, repeated measures analyses of variance (ANOVA) were conducted for all test sub-scores for Lesson 1 through Lesson 3. Overall analyses revealed main effects for Lesson on Control Score, $F(3, 138) = 77.50$, $p < .01$, Velocity Score, $F(3,138) = 21.07$, p

$< .01$, Speed Score, $F(3,138) = 62.44$, $p < .01$, and Points Score, $F(3, 138) = 114.99$, $p < .01$, and Total Score, $F(3, 138) = 138.66$, $p < .01$. No interactions were indicated. A mixed repeated measures ANOVA was also conducted for test game Total Scores for Lesson 1 through Lesson 3. A main effect for Total Score, $F(3, 138) = 139.08$, $p < .01$, was revealed. No interactions were indicated. Planned comparisons revealed no significant differences between groups for either Lesson 3 or the retention lesson, on either sub-scores or Total Scores.

Planned comparisons revealed no significant differences between groups for either Lesson 3 or the retention lesson, on either sub-scores or Total Scores. However, several sub-scores were in the predicted direction (see Table 2). Average control scores for Lesson 3 test games were higher for the massed group than for the spaced explicit group. This pattern held for retention test games. The trend for velocity sub-scores also favored the massed group over the spaced-explicit group both for Lesson 3 and retention games. The points sub-score was also in the predicted direction. Average points scores for Lesson 3 test games and retention games were greater for the spaced-explicit group than for the massed group. Not in the predicted direction were the speed sub-scores. Average Lesson 3 test game and average retention game speed scores were higher for the massed group than for the spaced-explicit group. Given the matched group design, a second, entirely within repeated measures analysis of variance (ANOVA) was conducted, using the massed participant as the first observation and the spaced-explicit participant as the second observation. Results revealed the same pattern of null results as the mixed design.

Learning Styles Questionnaire and Progress Report data await future, pooled analyses, as part of the larger research program. The larger sample size will enable correlational analyses to be conducted.

Discussion

No statistically significant effects of explicit interlesson processing or massed practice on sub-component scores were evidenced. However, trends were in the predicted direction for Control, Velocity, and Points sub-scores for both Lesson 3 test games and for retention games. The mean differences may be due to chance. Alternatively, the experiment may have failed to detect an actual difference, given its low power.

Two procedures may have decreased power by reducing effect sizes. One procedure was giving only two opportunities to elaborate between lessons. To increase the effect of explicit processing, more inter-lesson exercises may be required. The second procedure was including participants who used the strategy of wrapping the screen. "Wrappers" slowly fly their ship in either a horizontal or vertical plane, leaving the screen and reappearing at the opposite side of it. Including trainees that wrap reduces effect sizes under the assumption that these trainees are not attempting to learn the skill of flying the ship within the hexagon. According to the same assumption, these trainees reduce power further by adding inappropriate variability to the sub-scores. Experiment 3 addresses these issues.

Experiment 3

Attempts to increase treatment effects, reduce variance, and increase sample size were made in the present experiment. To increase the effect of interlesson processing, an explicit processing session was added between baseline games and Lesson 1 games. To

reduce variance, "wrappers" were excluded from an analysis comparing massed and spaced-filler groups. To increase sample size, subjects were pooled across studies.

In the present experiment, participants performed 3 Acquisition Lessons and 1 Retention Lesson on Space Fortress. Manipulated variables were spacing of practice, which were either 0 or 25 minute interlesson intervals, and opportunities for explicit processing, which were guided elaboration, problem-diagnosis, problem-solution, and mental rehearsal exercises between lessons or no explicit inter-lesson processing. The dependent variables were performance on sub-scores and Total Scores for test games.

The predictions are the same as those in Experiments 1 and 2. Control and Velocity should be improved by inter-lesson massing and a lack of inter-lesson explicit should be improved by inter-lesson spacing and inter-lesson explicit processing. The predictions for Control and Velocity are made with more confidence because they should be purer indicators of implicit processes, based on the assumption that higher skills in a hierarchy influence lower skills more than lower skills influence higher ones.

Method

Participants

The sample consisted of 66 male students from Texas A&M university who received credit in an Introductory Psychology course for their participation. Selection was based on an aiming screening task (Mane & Donchin, 1989) and a survey of video game experience. Subjects were not permitted to participate if they failed to obtain a minimum aiming score of 780 points or if they reported currently playing video games more than 20 hours per week.

Apparatus and Task

A laboratory room was equipped with four tables, each with an IBM PC/AT compatible computer, color monitor, joystick (Flightstick by CH Products), a three-button mouse, and two right-handed chair-desks. Two straps secured the joystick to the chair-desk and the subject's arm was supported comfortably in a horizontal position perpendicular to the joystick. The mouse was placed on the other chair-desk. The Space Fortress task used in the present study was a version of the game developed by Gopher et al. (1994). Space Fortress and the aiming screening task are described in Experiment 1. The math and keypress, the interlesson processing exercises, the Learning Styles Questionnaire, and the Progress Report are described in Experiment 2. The mathsolve/ keypress task was used for 2 reasons: (1) to make total time training equal for the massed group and (2) as a filler task to prevent inter-lesson explicit processing.

Experimental Design

Participants, assigned to one of three conditions, were matched on baseline Total Scores, using plus or minus 400 points as a criterion for matching. Condition served as a between-subjects independent variable. The conditions were: (1) massed (0 minute interlesson interval) (2) spaced-filler (25 minute interlesson interval, during which a filler task was performed) and, (3) spaced-explicit (25 minute interlesson interval, during which an inter-lesson processing exercise was performed). This between-subjects independent variable was factorially combined with 2 within-subjects independent variables: (a) Acquisition, Lessons 1-3 and (b) retention, Lesson 4. Time of day for training was counterbalanced across participants, who were equally and randomly assigned to morning, afternoon, or evening. The dependent variables were (a) sub-scores averaged over 2 test

games per lesson, (b) Total Score averaged over 2 test games per lesson, and (c) answers to questions in the Progress Report and Learning Styles Questionnaire.

Two separate matches were done, both using the criterion of plus or minus 400 points on baseline Total Score. The first matched massed subjects (n=22), spaced-filler subjects (n=22), and spaced-explicit subjects (n=22). The second matched massed (n=24) and spaced-filler (n=24) subjects who did not use the wrapping strategy, based on the previously stated criteria. Eliminating "wrappers" from the spaced-explicit group decreased the sample size too much to enable a match between all 3 conditions. Separate matches were deemed appropriate because elimination of wrappers was only critical for the comparison between the massed and spaced-filler groups.

Procedure

Subjects signed an informed-consent form, performed the aiming screening task, then filled out a survey of video-game experience. Subjects who passed these screening measures were randomly assigned to the two experimental conditions. Next, subjects watched a 17-minute instruction video before they played 4, 3-minute baseline games of Space Fortress. Following the baseline games, subjects in the massed condition had a 3 minute break before they watched a 5-minute "summary of instructions" video. Subjects in the spaced-filler and spaced-explicit conditions performed a filler task and inter-lesson processing exercises, respectively, before watching the summary of instructions. After the summary of instructions, all subjects performed 3 lessons of Space Fortress, each consisting of 8, 3-minute practice games and 2, 3-minute test games. (The Math Solve/ Keypress Task and the inter-lesson processing exercises were performed between Space Fortress Lessons 1 and 2 and between Lessons 2 and 3 by the filler and explicit groups,

respectively.) Next, all subjects filled out the Learning Styles Questionnaire. The interval between lesson 3 and the first retention lesson, during which subjects filled out the Learning Styles Questionnaire, was 3 minutes and was closely monitored. After the questionnaire, participants performed a retention lesson, which consisted of 2, 3-minute test games. Next, subjects completed the Progress Report. Following the Progress Report, subjects in the massed condition performed 3 sessions of the filler task. To conclude, subjects were debriefed.

Results

A mixed, repeated measures analyses of variance (ANOVA) was conducted for all test sub-scores for Baseline through Lesson 3 (see Figures 1-3). Overall analyses indicated main effects for Lesson on Control Score, $F(3, 138) = 62.75$, $p < .01$, Velocity Score, $F(3,138) = 21.20$, $p < .01$, Speed Score, $F(3,138) = 42.83$, $p < .01$, Points Score, $F(3, 138) = 83.79$, $p < .01$, and Total Score, $F(3, 138) = 138.66$, $p < .01$ and Lesson X Condition interactions for Control Score, $F(3, 138) = 4.01$, $p < .01$, Wraps Score, $F(3, 138) = 3.56$, $p < .05$, and Total Score, $F(3, 138) = 3.46$, $p < .05$.

One-tailed comparisons for the match between the spaced-filler and spaced-explicit groups revealed a main effect for Speed Score, $tF(1, 42) = 1.76$, $p < .05$. Comparisons also revealed significant differences on Speed Score for Lesson 1, $t(1, 42) = 2.34$, $p < .05$ and retention, $t(1, 42) = 1.87$, $p < .05$. Spaced-explicit scores were higher ($M = 159$; $SD = 169$) than spaced-filler scores ($M = 56$; $SD = 116$) for Lesson 1. Spaced-explicit scores were also higher ($M = 287$; $SD = 197$) than spaced-filler scores ($M = 182$; $SD = 171$) for retention.

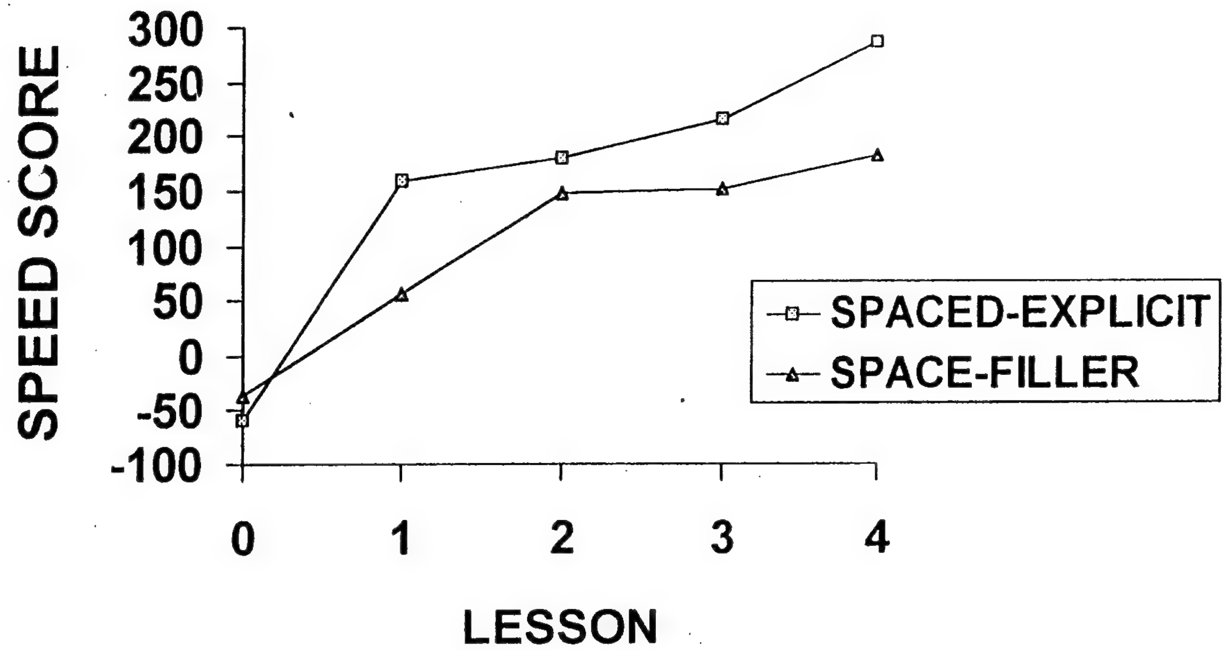


Figure 1. Average Speed Score of two test games on Space Fortress as a function of Lessons 1-4 for spaced-explicit and spaced-filler conditions

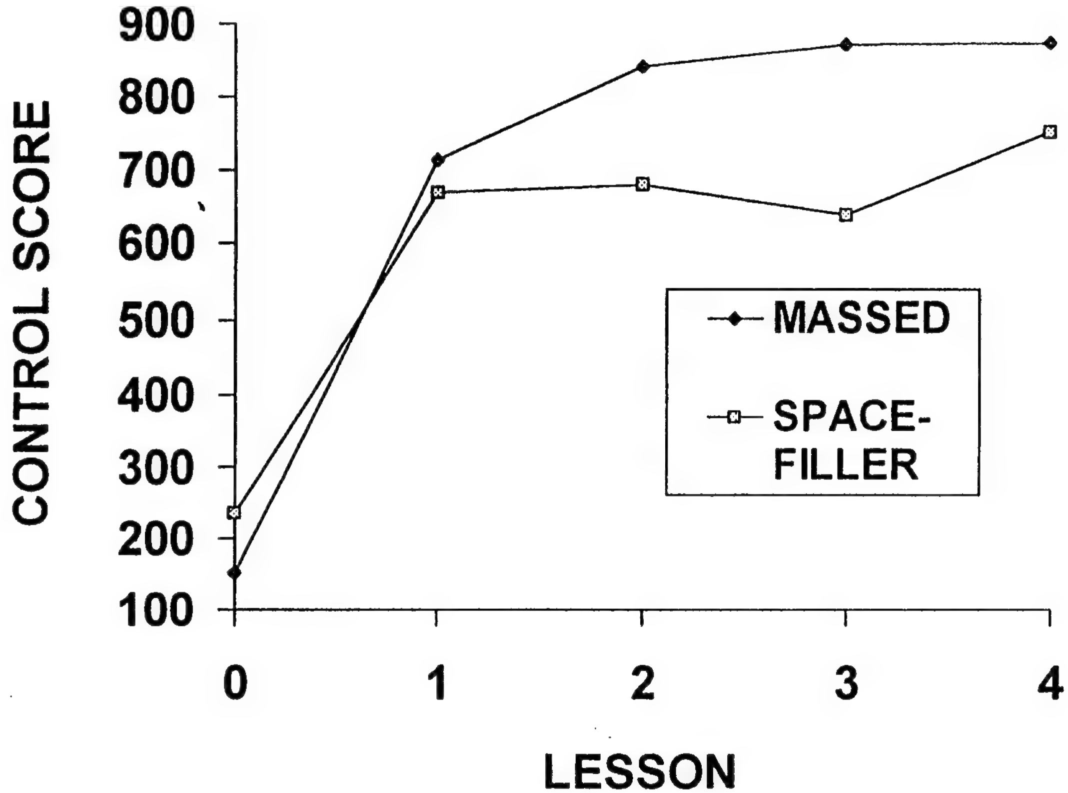


Figure2. Average Control Score of two test games on Space Fortress as a function of Lessons 1-4 for massed and spaced-filler conditions.

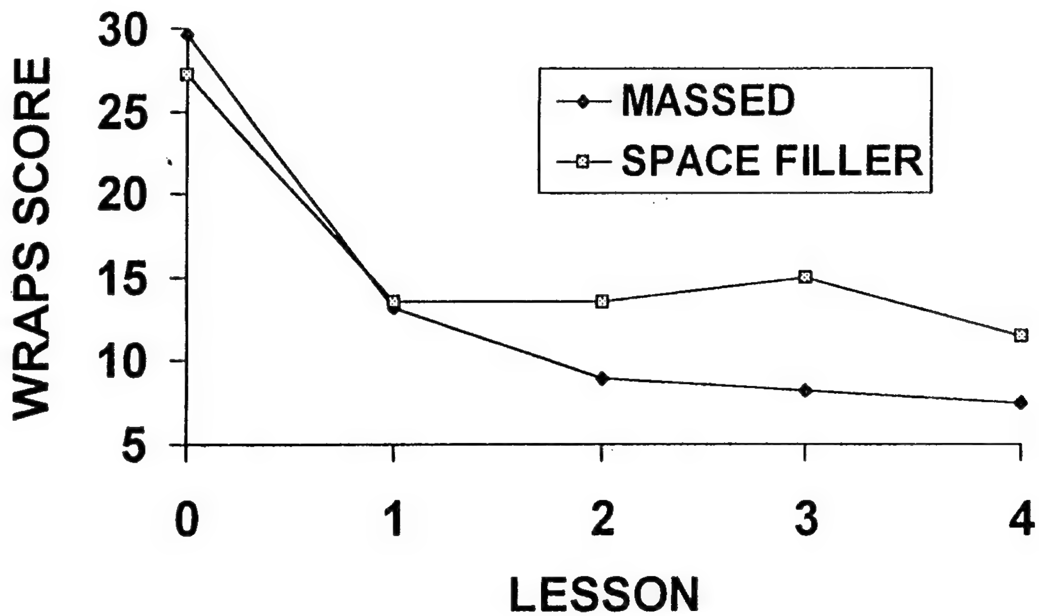


Figure 3. Average Wraps Score of two test games on Space Fortress as a function of Lessons 1-4 for massed and spaced-filler conditions.

For the match between massed and spaced-filler nonwrappers, one-tailed planned comparisons revealed significant differences on Control Scores for Lesson 2, $t(1, 46) = 1.71$, $p < .05$ and Lesson 3, $t(1, 46) = 2.46$, $p < .05$. Massed group Control Scores (Lesson 2, M = 841; SD = 255: Lesson 3, M = 871; SD = 253) were significantly higher than spaced-filler group Control Scores (Lesson 2, M = 681; SD = 380: Lesson 3, M = 639; SD = 385). The difference on Control Score approached significance for retention games, $t(1, 46) = 1.60$, $p = .055$, with massed Control Scores (M = 876; SD = 215) higher than spaced-filler Control Scores (M =754; SD = 302). Comparisons revealed significant differences between groups on Wraps Score for Lesson 2, $t(1, 46) = 1.75$, $p < .05$ and Lesson 3 test games, $t(1, 46) = 2.50$, $p < .01$. Massed condition wraps (Lesson 2, M = 8.88; SD = 7.14: Lesson 3, M = 8.12; SD = 7.03) were significantly less than spaced-filler wraps (Lesson 2, M = 13.5; SD = 10.79: Lesson 3, M = 14.90; SD = 11.29). A significant difference on Wraps Score remained for retention games, $t(1, 46) = 1.85$, $p < .05$, with massed condition wraps (M = 7.44; SD = 5.51) significantly less than spaced-filler wraps (M = 11.38 ; SD = 8.86).

Pooled analyses of Learning Styles and Progress Report will be analyzed in the future.

Discussion

The results suggest that highly massed practice lessons benefit implicit learning mechanisms in Space Fortress. A 0 minute interlesson interval produced better ship control than a 25 minute interval, as evidenced by Control and Wrap Scores. Wrap Score is another useful indicator of ship control, since participants that relied solely on a wrapping strategy, according to the criterion used, were eliminated from the analysis. No

benefit of massed practice was obtained for controlling ship velocity. However, velocity may not be the best indicator of learning ship control, since it is possible to obtain a high velocity score without learning the ship control required to circumnavigate the fortress.

Minimal support was obtained for a benefit of interlesson explicit processing on explicit learning mechanisms. Explicit processing between lessons produced higher Speed Scores than a 25 minute interval in which explicit processing was preempted. However, Speed Score is a relatively impure indicator of explicit learning, since the Speed sub-component is heavily influenced by skills that are learned implicitly, such as ship control. Furthermore, no support for a benefit of explicit on Points Score was obtained.

Results are inconclusiveness regarding the detrimental effects of massed lessons on explicit learning mechanisms and explicit inter-lesson processing on implicit mechanisms. Null results for the effect of massed lessons on Speed and Points sub-scores and and the effect of explicit processing on Control and Velocity sub-scores make additional research necessary.

In summary, the results indicate that massed practice improves learning of ship control, as indicated by Control Score and Wrap Score. This result suggests that implicit learning influences Control and Wrap Scores enough to produce an effect of massed practice lessons on these game components. The results also indicated that explicit inter-lesson processing improves learning of mine management, as indicated by the Speed Score. This result suggests that the influence of explicit learning on Speed Score is strong enough to produce a benefit of explicit inter-lesson processing for this subscore. However, this conclusion is made with less confidence, given that no difference was found for the final acquisition lesson.

General Discussion

These results have implications for three goals of the TRAIN laboratory. The goals are: (1) a cognitive taxonomy that decomposes tasks into psychologically meaningful knowledge/skill types, (2) a general learning model that specifies how those knowledge/skill types are acquired, and (3) a general instructional model that specifies how acquisition of those knowledge/skill types can be optimized. Towards a better cognitive taxonomy, the present study illustrates the importance of distinctions between explicit inter-lesson processing and no explicit inter-lesson processing and between implicit and explicit processing during complex skill acquisition. These distinctions are evaluated in the context of the componential iterative learning theory, which provides a framework for understanding implicit and explicit processing throughout complx skill development. Finally, the present results suggest that the infuence of interlesson intervals within the range of 25 min will have to be taken into account by a general instructional model . (Parts of this report were based on a draft of Kip Corrington's dissertation, which is being completed under my suppervision; more details will be in the final dissertation).

References

Ackerman, P. L. (1987). Individual differences in skill learning: An integratin of psychometric and information processing perspectives. Psychological Bulletin, 102, 3-27.

Corrington, K. A. (1994). Spacing effects in the acquisition and maintenance of a complex skill. Unpublished master's thesis, Texas A&M University, College Station, Texas.

Corrington, K. A. & Shebilske, W. L. (1995). Complex skill acquisition: Generalizing laboratoy-based principles to football. Applied Research in Coaching and Athletics Annual, 54-69.

Donchin, E. (1989). The learning strategies project: Introductory remarks. Acta Psychologica, 71, 1-15.

Gopher, D., Maya, W., & Bareket, T. (1994). Transfer of skill from a computer game trainer to flight. Human Factors, 36(3), 387-405.

Gopher, D., Weil, M., & Siegel, D. (1989). Practice under changing priorities: An approach to the training of complex skills. Acta Psychologica, 71, 147-177.

Lee, D. L., & Genovese, E. D. (1988). Distribution of practice in motor skill acquisition: Learning and performance effects reconsidered. Research Quarterly, 59 (4), 277-287.

Mane, A. M., & Donchin, E. (1989). The space fortress game. Acta Psychologica, 71, 17-22.

Reber, A. S. (1993). Implicit Learning and Tacit Knowledge. New York, New York: Clarendon.

Seger, C. A. (1994). Impliciting learning. Psychological Bulletin, 115 (2), 163-196

Shebilske, W. & Corrington, K. (1995). Spacing effects in complex acquisition and maintenance. (Submitted).

Shebilske, W. L., Geottl, B., & Regian, J. W. (in press). Individual and group protocols for training complex skills in laboratory and applied settings. In Gopher, D. & Koriat, A. (Eds.) Attention and Performance XVII: Cognitive regulation of performance: Interaction of theory and application. Hillsdale, NJ: Earlbaum.

Shebilske, W. L., Jordan, J. A., Arthur, W., & Regian, J. W. (1993). Combining a multiple emphasis on components protocol with small group protocols for training complex skills. Proceedings of the 37th Annual Meeting of the Human Factors and Ergonomics Society, 2, 1216-1220.

Shebilske, W. L., Regian, J. W., Arthur, W., & Jordan, J. A. (1992). A dyadic protocol for training complex skills. Human Factors, 34(3), 369-374.

TACTILE FEEDBACK FOR SIMULATION OF OBJECT SHAPE AND TEXTURAL INFORMATION IN HAPTIC DISPLAYS

Janet M. Weisenberger
Professor
Department of Speech and Hearing Science

Ohio State University
1070 Carmack Road
Columbus, OH 43210

Final Report for:
Summer Faculty Research Extension Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

December, 1995

TACTILE FEEDBACK FOR SIMULATION OF OBJECT SHAPE AND TEXTURAL INFORMATION IN HAPTIC DISPLAYS

Janet M. Weisenberger
Professor
Department of Speech and Hearing Science
Ohio State University

Abstract

Two experiments were performed to follow studies completed as part of the 1994 Summer Faculty Fellowship Program. This line of research has both basic and applied motivations. The basic science questions addressed by these investigations concern how the human sense of touch, or tactile system, processes complex stimuli to arrive at percepts of object shape and surface texture. The applied aspects of this work involve the evaluation of tactile feedback provided by displays that can be actively moved by an observer (i.e., haptics). The goals of this work are to increase our understanding of the mechanisms underlying the perception of object shape and surface texture, and to improve designs for tactile feedback displays to be used in teleoperation and virtual reality applications.

The specific experiments examined 1) how changes in the size of a tactile display (the tactile "field of view") interact with the size and spatial frequency content of complex virtual patterns to impact the resulting percept; and 2) whether observers can discriminate among sets of shapes that contain different vibratory frequencies for their outer borders (edges) and inner detail. Although performance levels for different stimulus sizes were similar except for very small stimuli, the amount of time necessary for stimulus identification was affected by stimulus size. Results suggest that a large tactile field of view may not be absolutely necessary for accurate performance, but the ease with which the task can be performed will be greater for larger displays. The second experiment defined boundary conditions for using vibratory frequency and perceived intensity to represent different object properties. The results of these experiments, taken as a whole, will provide useful information for improvements in the design of tactile displays, and will begin to provide a metric of stimulus dimensions that could be used in mapping the features of complex objects in realistic sensing.

Introduction and Background

The long-term goals of this project are to investigate aspects of the tactile perception of surface shape and texture, for improvements in the design and implementation of tactile displays for use in teleoperation and virtual reality applications. A more thorough understanding of how observers process complex tactile stimuli will provide insights into how best to design tactile interfaces for use in such applications. Relevant areas of study include tactile psychophysics (basic detection and discrimination of tactile stimuli that vary in spatial and temporal characteristics), tactile pattern perception, and haptics

(kinesthetic feedback from active touch).

Under most circumstances, normal human sensing of object properties of shape and surface texture is accomplished via active movement of the skin surface (typically the fingertips) along the object to be sensed. This motion appears to be essential for accurate sensing of object properties. In fact, as Katz (1989) has pointed out (see also Lederman & Pawluk, 1993), in the absence of relative motion between the skin and the object, perception is reduced to simple binary discriminations. Activity among the tactile mechanoreceptor populations is also quite different when motion is not present. Lederman and Pawluk note that without such scanning motion, the contributions of rapidly-adapting (RA) and Pacinian (PC) mechanoreceptors are eliminated from the neural code, and the activity of slowly-adapting (SA) mechanoreceptors is substantially reduced compared to levels observed when scanning occurs (Johnson & Lamb, 1981). Whereas Connor et al. (1990) have found that a neural code based on the variance in firing rate of SA I mechanoreceptors correlates well with perceptual acuity in texture discriminations, data from Johansson and Westling (1987) suggest that RA afferents may mediate perception of surface textures too fine to be resolved by the SA I afferents. Results such as these suggest a role for both mechanoreceptor systems in the normal perception of object characteristics.

Nonetheless, a considerable literature has arisen from studies of the tactile system's ability to resolve stimuli presented in a passive mode to a stationary skin surface. The motivation for this line of research has been twofold. One driving force has been the desire to measure the spatial, temporal, and intensitive data-handling characteristics of the tactile system under stimulus conditions analogous to those employed in studies of the visual and auditory modalities. Another impetus has been the development of devices to use the tactile system as an alternative communication channel for persons with auditory or visual impairments. Such devices, which transmit speech or printed information via spatiotemporal tactile patterns to substitute for the impaired modality, are typically placed on the fingertip or other sensitive body area. Perception of the patterns presented on these devices does not require active movement by the observer.

Much of the work in identification of complex tactile patterns presented on stationary displays has been performed by Craig and his colleagues (Craig, 1976, 1977, 1978, 1979, 1980, 1981, 1982, 1983, 1985; Craig & Evans, 1987; Evans & Craig, 1986). Craig's work has employed the two-dimensional vibratory display of the Optacon (Telesensory Systems), a reading aid for blind persons. The display consists of 144 piezoelectric stimulators, vibrating at a frequency of 230 Hz, that contact the index fingertip of the user. In studies of pattern recognition with this display, Craig has mapped the time course and spatial extent of recognition masking and examined attentional determinants of tactile processing. Although the finger is placed on a stationary display, movement of tactile patterns can be simulated by spatiotemporal pattern variation (e.g., "scanning" or "drawing" movements).

Perhaps one of the most surprising findings of these pattern recognition studies, in light of the

abovementioned importance of movement between skin and stimulus, was Craig's (1981) report that pattern identification accuracy was better for a "static" stimulus (no simulated movement) than for any of four presentation modes that employed simulated movement: scan, slit-scan, continuous-sequential, and discontinuous-sequential drawing modes. It might have been expected that the introduction of a movement cue would bring the stimulus situation closer to that typically encountered in the real world, and thus that accuracy would be enhanced.

In Craig's 1981 study, stimuli were presented over a range of durations from 4 to 1000 ms. At long durations (greater than 400 ms), most of the modes of presentation showed similar levels of performance. As duration decreased, however, performance differences emerged. The static mode was relatively unaffected by duration down to 25 msec, below which performance levels dropped markedly. In the scan mode, reducing duration below 400 ms resulted in a systematic, and substantial, drop in performance. Performance in the slit-scan, continuous-sequential, and discontinuous-sequential modes also decreased at shorter durations.

The results were in contrast to earlier findings of Loomis (1974), who presented patterns on the tactile display of the TVSS, a visual substitution device for blind persons that converted optical stimuli to patterns of vibration on a 20 x 20 display worn on the back (White, Saunders, Scadden, Bach-y-Rita, & Collins, 1970). Four modes of presentation were employed: static, scan, and two slit-scan modes. Loomis found that the static and passive scan modes yielded much worse performance than the two slit-scan modes, and argued that the tactile system performed poorly when forced to rely only on spatial frequency information for pattern identification. Loomis proposed that the two slit-scan modes provided phase information in addition to spatial information, and thus produced better performance. Earlier results from Beauchamp, Matheson, and Scadden (1971) supported this notion. However, in Craig's 1981 study, the slit-scan mode actually yielded the worst performance.

Two possible differences might account for the results in the two studies. First, data from either study may have been device-specific, a question addressed below. Second, Loomis' stimuli were much longer in duration, exceeding 1 s. At these durations Craig found static and scan modes to be comparable, although the slit-scan mode still produced poorer performance even at this duration. An additional complication is the report of Loomis (1980), who used the Optacon display in a study of pattern perception with stimuli of different sizes, and found that small patterns were more recognizable when the slit-scan presentation mode was used, but that the static mode produced better performance for larger letters. In this study, durations were again long (1.5 s) relative to Craig's durations. However, the use of the Optacon display in Loomis' 1980 study suggests that the results are not specific to a particular device.

Even so, it is important to consider the activity of the various mechanoreceptor populations in responding to stimuli presented on these displays. Results from Gardner and Palmer (1989) suggest that the 230-Hz Optacon stimulus excites primarily the PC and RA afferent fibers, with relatively little activity

observed in the SA afferents. Thus it is possible that stimulus movement is not necessary for this task, given that RA and PC fibers are already active in the 230-Hz static presentation mode. However, this does not explain why identification performance in Craig's study was poorer when movement cues were introduced.

Another possibility to consider is that the simulated movement cues may not have been good approximations of active, or haptic movement by the human observer. Sherrick and Rogers (1968) and Kirman (1974) reported interstimulus element intervals of 50-100 ms for optimal perception of movement. In Craig's moving stimulus presentation modes the intervals between successive stimulus elements were much shorter at brief total stimulus durations. It is possible that the stimulus perceived with real haptic movement is quite different from that presented under a passive scan condition.

However, the role of the "haptic" aspect of the stimulus movement is also not clear. Both Lederman (1981) and Lamb (1983) found in texture discrimination studies that there was no difference in performance when active movement of the finger over the surface was permitted, compared to passive movement of the surface under a stationary finger, as long as contact force was controlled. Although differences in mechanoreceptor activity in the active and passive movement presentation modes would not necessarily be anticipated, nonetheless it would seem that active control of the movement would provide additional kinesthetic or motor cues to facilitate perception.

It may be that the perception of vibratory pattern stimuli is quite a different task from the perception of surface texture. It may also be that simulated movement is perceptually very different from actual movement for vibratory pattern identification. These questions are of some interest not only for a basic understanding of the processing characteristics of the tactile system, but also for the development of tactile stimulator arrays designed to simulate real experiences, or to present encoded information for supplementing vision or hearing.

The relative effects of active and simulated stimulus movement on identification of tactile patterns were investigated by Weisenberger and Hasser (1994) using a tactile array designed for telerobotic and virtual environment applications. This display differed from that of the Optacon used by Craig and Loomis in several respects. First, the Optacon stimulators were piezoelectric bimorphs with a very narrow stimulus bandwidth, whereas the stimulators in the display used in the present study were a titanium-nickel shape-memory alloy (SMA) with a broader bandwidth. These were activated by superimposing a pulse-width modulated current on an initial DC. Unlike the Optacon's fixed-frequency 230 Hz signal, frequency and duty cycle of the SMA display's stimulus, as well as the duration of the initial DC component, can be varied in software. This flexibility permits evaluation of pattern presentation conditions in which the SA mechanoreceptors might play a dominant role (i.e, by using low vibration frequencies).

A second difference between the Optacon and the SMA display is the density of stimulator elements. The original Optacon contained 144 pins (6 columns x 24 rows), covering approximately 1.1 x

2.7 cm on the fingertip. The SMA display had 30 pins (5 cols x 6 rows), covering a similar area. With an interpin distance of approximately 3 mm, the SMA display is much less dense than the Optacon display.

Finally, and most importantly for addressing the role of stimulus movement, the SMA display has been mounted on the pointer of a digitizing pad, and "virtual" patterns can be constructed in software such that active scanning movements of the display across the digitizing pad can encounter a pattern and move across it in any direction. This aspect of the apparatus permits the inclusion of an active, or haptic, movement presentation mode in the pattern identification task. While attempts have been made to mount an Optacon display on a microswitch assembly (Kuc, 1990) or a mouse (Boyd et al., 1990), pattern recognition per se was not the focus of these studies, and comparisons to static and passive scan presentation modes were not performed. A display more similar to the present work was designed by Garland (1974), who constructed an experimental one-handed Optacon. This Optacon incorporated both camera and tactile display in a single unit, as opposed to the standard Optacon, in which the hand-held camera was used to scan text, which was then presented to a non-moving tactile display. Garland's one-handed Optacon could be moved across text in a truly "haptic" mode. Although Garland reported encouraging reading rates for this device, further work with the display was not pursued.

In Weisenberger and Hasser's study, identification of a set of pattern stimuli was compared for three modes of presentation: static, passive scan, and haptic scan. Vibration frequency for the pattern stimuli was varied, to provide conditions in which all three tactile mechanoreceptor populations (SA, RA, PC) were activated. Results indicated that the haptic mode produced superior performance when a set of relatively complex patterns (letters of the alphabet) was employed. Weisenberger and Hasser cited a number of possible factors that may have contributed to the superiority of the haptic mode. These included the fact that under the haptic mode, observers were unconstrained in their behavior, and could scan the stimulus pattern as often as desired before responding. Further, they could scan the pattern in any direction they wished. In addition, total stimulus duration, given the repeated scanning, was longer for the haptic mode.

In followup experiments conducted as part of the 1994 Summer Faculty Research Fellowship, two possible explanatory factors were addressed. In one experiment, observers were permitted to repeat the stimulus as desired under three modes of stimulus presentation: static, passive scan, and haptic scan. In a second experiment, observers were permitted, under the passive scan mode, to present the stimulus from multiple directions, to determine whether the presentation of different information in stimulus repetitions led to higher levels of performance than simple repetition of the same information.

Results of these experiments indicated that the ability to repeat the stimulus accounted for much of the superiority of the haptic mode observed by Weisenberger and Hasser (1994). However, in the second experiment, the condition in which each stimulus repetition was presented from a different scanning direction was superior overall to the condition in which stimulus repetitions all had the same

direction. These results suggested that in repetition of the stimulus, advantage is gained from presenting different information on subsequent scans, as compared to repeated presentations of the same information. It was also noted that some scan directions provided more information than others, as evidenced by considerable differences in performance for different directions.

A second series of experiments during the 1994 Summer Faculty Fellowship began investigation of a related question in tactile pattern perception. In these studies, the impact of reductions in the tactile "field of view" was evaluated. Loomis (1980), as mentioned, found that a reduced-field slit-scan mode yielded higher levels of performance than the full-field mode when stimulus patterns were small and spatially dense. Loomis (1981) argued that the superiority of sequential presentation of stimuli might result from the limited spatial resolution of the tactile sense. If the tactile system acts as a low-pass filter for spatial sensing (relative to vision, for example), then a sequential scan of adjacent elements would give rise to a distribution of relative intensity of stimulation across the dimensions of the pattern. A simultaneous presentation of all elements in the pattern would yield a less accurate representation since the intensity of that pattern at each point is the convolution of the point-spread function at a given point with the entire input pattern, rather than the product of the pattern and the scanning aperture.

However, Loomis noted that the superiority of reduced-field presentation modes would be expected only for situations in which the spatial resolution abilities of the tactile system were challenged, i.e., for patterns complex enough to push the low-pass properties of the system. If so, then it must be assumed that the findings of a number of other studies (e.g., Craig, 1981), indicating that the sequential and slit-scan presentation modes were actually inferior to a full-field presentation, must have employed stimuli that did not challenge the spatial resolution abilities of the tactile system.

Results supporting Craig's (1981) finding were reported in Hill's (1974) study, in an additional experiment in which the number of columns of stimulation on the Optacon display was manipulated. On one fingertip, subjects were tested with 2, 4, 6, or 8 columns of simultaneous activation, displaying successively greater portions of an alphabetic pattern. Hill found a systematic increase in letter recognition accuracy as the number of columns was expanded from two to eight, suggesting that a larger field of view resulted in improved pattern recognition.

However, all of the above results were obtained with stationary tactile displays. That is, even though some of the presentation modes simulated pattern movement, the finger itself did not move. This is quite different from normal human sensing of object and surface properties, in which information is enhanced by movement of the finger relative to the surface being sensed.

Given the possibility that pattern perception with active movement of the fingertip, under the control of the observer, might be qualitatively different in some respects from pattern recognition with stationary displays, the question arises whether reductions in the tactile field of view would yield deleterious effects similar to those observed by Craig (1981) and Hill (1974). To the degree that spatial

extent is reduced on the display, the temporal aspects of the stimulus might be emphasized. The spatial aspects of the stimulus would be increasingly dependent on the spatial feedback provided by kinesthetic cues from active movement. In the limit, if the field of view is reduced to a single element, the temporal aspects of the display are maximized.

There are a number of real-world tasks in which observers perform with a reduced tactile field of view. Katz (1925; 1989 translation) reported experiments in which individuals performed tactile tasks using an instrument to sense the surface, rather than by direct contact between skin and surface. In one study, subjects were able to identify the hardness of pencil leads by moving them across a piece of paper, based on vibrational cues transmitted to the fingers from friction between the pencil lead and the paper. In another study, subjects were able to discriminate the roughness of different grades of paper by moving a wooden rod across their surfaces. In these studies, surface features were sensed in an essentially temporal manner, since contact with the surface was a single point. Spatial cues were generated only kinesthetically. Results such as these suggest that when haptic movement is permitted, a reduced field of view might not be deleterious.

There is a strong, practically-motivated impetus for designing small and efficient tactile displays for use in virtual reality, telerobotic, and sensory aids applications. Displays with a smaller number of transducers would most likely be smaller in size, and might also impose less formidable processing and memory requirements, than larger displays. If haptic movement were a characteristic of the display situation, it is possible that a full-field display of actuators would not be required for accurate object and surface sensing. Results such as those of Katz support such a notion. However, the task used by Katz is very different from pattern identification.

In the experiments conducted during the summer of 1994, the question of whether haptic scanning of vibratory patterns is impaired by reductions in field of view was empirically addressed. One experiment employed a single-fingertip square display consisting of 9 elements, covering an area of 0.6 cm^2. Patterns used in the experiment were identical to those from an earlier study in which a full-fingertip, 30-element display (1.2 x 1.5cm) had been tested. The pattern sizes were chosen to fill this full-fingertip display, and thus were larger than the 9-element display used in the first experiment of the present study. Thus, the 9-element display constituted a reduced field of view for these patterns. The experiment was conducted to determine whether these patterns could be identified with a reduced field-of-view display. Results indicated that reducing the display size from 30 to 9 elements had no effect on performance, suggesting that considerable reduction in display size can be implemented without deleterious effects. A further experiment attempted to determine just how far the field of view could be reduced before performance was impaired. Display sizes of 9, 4, and 1 element were compared. Results showed that reduction to 4 elements had only a slight effect on performance, but that reducing display size to a single element, which maximized the temporal aspects and minimized the spatial aspects of the display, caused a

serious decrease in performance levels. For a set of letter stimuli, performance for the 1-element display fell to nearly chance levels.

In the present study, two lines of experimentation were begun. In the first, the field of view work described above was extended. In the second, we began to address how physical aspects of the stimulus delivered to the vibratory display (such as frequency, duty cycle, etc.) can be used to facilitate discrimination of surface textural features.

The experiments addressing field of view attempted to elucidate further the relationship between character size and field of view, by creating very small and very large stimulus sets, to be sensed by different display sizes. If Loomis's argument regarding the superiority of small displays for very complex patterns is correct, then it is possible that the patterns used in the experiments described above did not sufficiently tax the spatial filtering characteristics of the tactile system. Four different tactile displays were tested: 2, 30-element displays, with different overall dimensions due to different interelement spacing (1.5 vs 3 mm); a 24-element display with 1.5 mm interelement spacing; and a 9-element display with 3-mm interelement spacing.

In the second experiment, subjects identified virtual objects that were generated with different vibratory frequencies for border lines and interior detail (e.g., a shape that has a low-frequency, high-amplitude border lines but high-frequency, low-amplitude interior detail, as compared to one that has the opposite features).

Methodology for each of these experiments is described in the following section.

Experiment 1 -- Interactions of character size and tactile field of view

Methods

Subjects. A total of four subjects served in each of the proposed experiments. Subjects were 3 female and 1 male adults with normal hearing, normal vision, normal tactile detection thresholds, and no report of tactile dysfunction. All subjects were practiced extensively with the stimulus conditions to be used in the present experiment prior to the onset of actual data collection.

Apparatus. The shape-memory alloy (SMA) and dot-matrix displays used in the present study were designed and constructed by the TiNi Alloy Corporation (A. Johnson, 1991). Four different displays were tested. Display 1 contained 30 element arranged in 5 columns by 6 rows, with an interelement distance of 3 mm. Only the top 5 rows of the display were actually used for generating stimuli, such that the overall stimulus dimensions were 1.2cm^2. Display 2 contained 9 elements in a 3x3 arrangement, with an interelement distance of 3 mm, and overall dimensions of 6 mm^2. This display was in effect a reduced field of view from the first display (although the actuators themselves had a slightly different design). Display 3 contained 30 elements, again in a 5x6 arrangement. However, this display had an interelement distance of 1.5 mm, and utilization of the top 5 rows produced overall display dimensions of 6 mm^2. This display was used because it had the same number of elements as Display 1, but the same overall

dimensions as Display 2. Display 4 contained 24 elements in a 4-column by 6-row arrangement, and interelement distance of 1.5 mm. Using the top 5 rows produced overall dimensions of 4.5x6 mm. Although this display was functionally a reduced field of view for Display 3, it utilized a very different actuator technology (dot-matrix printhead rather than shape-memory alloy) ans was included primarily for this reason. All displays were mounted in a plastic casing on which the fingertips could be rested. The stimulator elements, when activated, protrude through holes in the casing to contact the skin surface.

The actuators for the SMA displays are constructed of a metal rod, connected to a thin length of titanium-nickel shape-memory alloy (SMA) wire. When heated by input current, the SMA wire contracts, forcing the rod upward; when current is discontinued, the wire cools and returns to its original configuration. Use of an input pulse-width-modulated current results in a vibratory output. Variations in pulse rate, frequency, duty cycle, and stimulus duration, are controllable in software. A more extensive characterization of a similar display can be found in Hasser and Weisenberger (1993).

A CalComp Drawing Board II digitizing pad was employed to permit active movement of the display. The SMA display is mounted atop the pointer for the pad, and can be moved about the pad surface with preservation of the X-Y coordinates of the movement. The digitizing pad communicates with the computer via the mouse port.

The display and digitizing pad are interfaced to an 80386-based PC computer. The PC's serial port is used for communication with the display. C software permits specification of experimental runs, including values for each of the variables mentioned above, and controls stimulus presentation, response collection, and data storage. For Experiment 1, vibratory frequency was fixed at 31-33 Hz, and duty cycle was manipulated for each display to achieve comparable perceived loudness across displays. Stimulus duration was controlled by the subject by the total duration of active scanning movements.

Two stimulus sets were employed. Each set contained the same patterns (for consistency, we employed the set of letters of the alphabet used by Weisenberger and Hasser (1994) and in the experiments conducted as part of the 1994 Summer Faculty Fellowship program). These stimuli were approximately 1.2cm^2 in size, and included the letters A,K,X,B,R,G,C,D,O,Q. The difference across sets for Experiment 1 was stimulus size; stimulus sizes that are two and four times the size of the previous stimuli, as well as one-half the size, will be employed. For comparison purposes, the original stimulus set was employed as a control condition. It is important to note that the original stimulus set was constructed to fill 5 columns by 5 rows of the display. One question of interest was how to display these stimuli on the smaller and denser 30-element display. It was decided to reduce the overall dimensions of the stimuli to fit the smaller display, such that the overall dimensions were now 4 mm^2, rather than the 12 mm^2 of the larger display. For the remaining stimulus sets, which were one-half, 2, and 4 times as large, corresponding reductions were employed when the smaller 30-element and 24-element displays were tested. While this limited direct comparability based on stimulus set size, the comparison was of interest in determining

whether more densely-packed stimulus features were equally identifiable.

All stimulus presentations were in the haptic scan mode, in which the pattern is constructed as a set of points at a particular location on the digitizing tablet in X-Y space, which causes activation when any portion of the SMA display reaches the location of any point in the set. Movement of the digitizing tablet's pointer, with the SMA display mounted atop it, across the pad gives rise to the percept of a pattern in a fixed location on the pad.

Procedure. Testing was conducted in a quiet area of the laboratory, with the subject seated approximately 0.5 m from the PC monitor. The left index finger of the subject rested lightly on the plastic enclosure of the display, and responses were entered on the keyboard of the PC with the right hand.

Stimulus presentation was arranged in 40-trial blocks. On each trial, a stimulus was presented, and subjects were prompted to choose the number corresponding to the icon of the presented pattern. Icons for each of the 10 patterns were displayed on the monitor throughout the block of trials. After a response was entered, trial-by-trial feedback was provided on the monitor. In the haptic scan mode, subjects were permitted to scan in any direction for as long as desired before entering a response.

Within a 40-trial block, patterns were presented in a random order. Testing was conducted in sessions lasting 1-2 hours, with frequent rest periods to minimize fatigue.

Results and Discussion

Initial data were collected with a set of patterns that had been shown to produce near-perfect performance levels inpast studies. These patterns consisted of single-line or two-line segments arranged in different horizontal, vertical, or diagonal positions. These patterns were tested in the present study to determine comparability of displays (I.e., to ensure that patterns that were very easy to discriminate on one display were equally easy to discriminate on other displays. Results showed performance levels averaging 94% correct across all displays, suggesting that different display sizes and technologies were not a factor when patterns were easy to discriminate. However, the near-ceiling performance levels do not predict how the different displays will transmit more complex patterns.

For the set of letters, such high performance levels were not found. However, previous work with this same pattern set showed average performance levels between 60 and 65% correct for the standard-size (1.2 cm^2) set. Thus, such high levels were not anticipated in the present work. The results for the letter stimuli are shown in Figures 2 and 3, and show percent correct for each display, averaged across subjects and sessions, as a function of the size of the patterns in a set.

In Figure 2, performance for the larger 30-element and 9-element displays is shown. Each of these displays was used for sensing the same pattern sizes, and thus results are shown together. As can be seen, there is very little difference in performance for the 9-element and 30-element displays, suggesting that this degree of reduction in field of view has little effect on perception. Of further interest is the fact that the size of the stimulus did not affect performance either, at least for stimuli between 12 and

48 mm^2 in size. Performance did drop substantially for the smallest 6mm stimuli, but did so for both displays. In Figure 3, performance for the remaining two displays, those with the denser spacing, is shown. Overall, for the larger pattern sizes, performance is even better for the smaller 30-element display than for the larger displays. Performance with the 24-element dot matrix display is comparable to that for the larger displays. Again, as noted in theprevious figure, there is a substantial drop in performance for the smallest pattern sizes (in this case 3 mm^2). Although still above chance levels, performance for these very small patterns is not impressive.

Overall, these results suggest that pattern size and display size do not interact in a signficant way for patterns above a certain size (I.e., 6 mm^2). Pattern size becomes a significant issue only when patterns are very small relative to the size of the display. This may have more to do with the resolution of the display than with the absolute size of the patterns. That is, it was not possible to makethe patterns any smaller for the large 30-element display than we did, because display resolution limited reduction ability. It would be of some interest to explore further subjects' abilities to identify very small patterns. In addition, anecdotal reports from the subjects indicated that they did not like the dot-matrix display in this task, because they felt that display intensity was high relative to the other displays, and that receptor fatigue occurred more rapidly. This issue also merits further study.

Experiment 2 -- Haptic sensing of objects

Method

Subjects. Subject selection details were as in Experiment 1.

Apparatus. General apparatus details were as in Experiment 1, except that only one display was used (the reduced-size 30-element SMA display , Display 3, from Experiment 1). Stimuli for Experiment 2 consisted of two-dimensional objects, generated with different vibratory frequencies and different duty cycles for border lines and inner detail. For example, patterns might have border lines presented at 20 Hz, and a 50% duty cycle, but have inner detail presented at 200 Hz and 25% duty cycle. Frequency varied between 5 and 200 Hz, and duty cycle between 10 and 80%. These ranges of stimulation should produce several identifiable steps within each stimulus dimension, and possibly more values when dimensions are combined. Stimuli were combined into sets of six items (higher-frequency border, blank interior; higher-frequency border, higher frequency interior; higher-frequency border, lower frequency interior; lower frequency border, blank interior; lower frequency border, lower frequency interior; and lower frequency border, higher frequency interior). Subjects were asked to identify which stimulus from the set was presented on each trial.

Procedure. General procedural details were as in Experiment 1.

Results and Discussion

Preliminary results for all subjects are shown in Figure 4, for the case where an amplitude cue was

present. In all cases, the higher frequency stimulus part had a duty cycle of 30%, and the lower frequency stimulus part had a duty cycle of 80%. Each pair of frequencies tested is listed on the abscissa. As can be seen, performance is best when there is a large difference between the high and low frequencies, and falls systematically as the frequency separation between the border and inner area of the rectangle is reduced. However, even at the smallest frequency separation tested (50/100 Hz) performance is still above chance levels (chance was approximately 18 percent in this task). These data suggest that subjects do have the ability to distinguish a low-frequency cue from a high-frequency cue when an amplitude cue is present.

Figure 5 shows the data for the case when no amplitude cue was present. It can be seen that the use of an amplitude cue did not affect performance when the frequency separation was large, i.e., for the 05/100 and 10/100 Hz cases. However, performance for smaller frequency separations was much lower when the amplitude cue was not present, as compared to the performance shown in Figure 4.

Overall, these data suggest that vibratory frequency is a potentially useful cue for distinguishing the exterior edges and interior details of a surface, particularly when an amplitude cue is also present (as would typically be the case in normal sensing. Future studies will continue this line of investigation by evaluating additional combinations of frequency and other amplitude cues, to establish boundary conditions and determine the limits of information transmission for these combinations of cues.

Conclusions and Implications

The results of the proposed studies addressed the two goals of our overall research project: to advance our knowledge of how the human tactile system processes the complex aspects of objects encountered in the environment, such as surface shape and texture, and to improve the design of tactile displays for virtual sensing of objects. The first experiment specifically provided data relevant to the question of how far the tactile field of view can be reduced without degrading perception, a question that is of interest on both basic and applied levels. As a basic science issue, it is of interest to determine how the spatial aspects of stimulus sensing can be preserved when the stimulus is presented sequentially, a portion at a time, such that spatial and temporal information must be integrated by the observer. The applied implications of this work are obvious, in that tactile displays are at present large and cumbersome to produce, and the ability to reduce display size would be of great use in improving existing designs, particularly for fitting tactile displays into a larger apparatus for multimodal virtual sensing.

The second experiment determined whether observers can discriminate among objects that differ in the tactile characteristics of their outer borders and inner detail. The results of this experiment provide further insights into how multidimensional stimulus coding can be used in realistic representation of object features. Because knowledge of the perceptual properties of the tactile system is not as advanced as that for other sensory modalities, such as vision and audition, and because the development of displays

for virtual sensing is much less advanced for the sense of touch than for vision or audition, considerable research effort, both basic and applied, is necessary for the development of tactile displays that match the sophistication of currently -available visual and auditory displays.

BIBLIOGRAPHY

Beauchamp, K.L., Matheson, D.W., and Scadden, L.A. (1971). Effects of stimulus-change method on tactile-image recognition. Perceptual and Motor Skills, 33, 1067-1070.

Boyd, L.H., Boyd, W.L. and Vanderheiden, G.C. (1990). The graphical user interface crisis: Danger and opportunity. No. September, Trace R&D Center, University of Wisconsin-Madison.

Connor, C.E., Hsiao, S.S., Phillips, J.R. and Johnson, K.O. (1990). Tactile roughness: Neural codes that account for psychophysical magnitude estimates. Journal of Neuroscience, 10, 3823-3836.

Craig, J.C. (1976). Vibrotactile letter recognition: The effects of a masking stimulus. Perception & Psychophysics, 20, 317-326.

Craig, J.C. (1977) Vibrotactile pattern perception: Extraordinary observers. Science, 196, 450-452.

Craig, J.C. (1978). Vibrotactile pattern recognition and masking. In: Active Touch--The Mechanisms of Recognition of Objects by Manipulation: A Multidisciplinary Approach. Ed. G. Gordon (Oxford: Pergamon Press).

Craig, J.C. (1979). A confusion matrix for tactually presented letters. Perception & Psychophysics, 15, 123-127.

Craig, J.C. (1980). Modes of vibrotactile pattern generation. Journal of Experimental Psychology: Human Perception and Performance, 6, 151-166.

Craig, J.C. (1981). Tactile letter recognition: Pattern duration and modes of pattern generation. Perception & Psychophysics, 31, 540-546.

Craig, J.C. (1982). Vibrotactile masking: A comparison of energy and pattern maskers. Perception & Psychophysics, 31, 523-529.

Craig, J.C. (1983). The role of onset in the perception of sequentially presented vibrotactile patterns. Perception & Psychophysics, 34, 421-432.

Craig, J.C. (1985). Attending to two fingers: Two hands are better than one. Perception & Psychophysics, 38, 496-511.

Craig, J.C. and Evans, P.M. (1987). Vibrotactile masking and the persistence of tactual features. Perception & Psychophysics, 42, 309-317.

Evans, P.M. and Craig, J.C. (1986). Temporal integration and vibrotactile backward masking. Journal of Experimental Psychology: Human Perception & Performance, 12, 160-168.

Gardner, E.P. and Palmer, C. I. (1989c). Simulation of motion on the skin: II. Cutaneous mechanoreceptor coding of the width and texture of bar patterns displaced across the OPTACON. Journal of Neurophysiology, 62, 1437-1460.

Garland, H. (1974). An experimental one-handed Optacon. American Foundation for the Blind Research Bulletin, 28, 165-168.

Hasser, C.J. and Weisenberger, J.M. (1993). Preliminary evaluation of a shape-memory-alloy tactile

feedback display. ASME Advances in Robotics, Mechatronics, and Haptic Interfaces, DSC-49, 73-80.

Hill, J.W. (1974). Limited field of view in reading letterhapes with the finger. In F.A. Geldard (Ed.), Cutaneous communication systems and devices. Austin, TX: The Psychonomic Society.

Johansson, R.S. and Westling, G. (1987). Signals in tactile afferents from the fingers eliciting adaptive motor responses during precision grip. Experimental Brain Research, 66, 141-154.

Johnson, A.D. (1990). Shape-memory alloy tactile feedback actuator. Technical Report AAMRL- TR-90-039. TiNi Alloy Company, 1144 65th Street, Oakland CA, August.

Johnson, K.O. and Lamb, G. D. (1981). Neural mechanisms of spatial tactile discrimination: neural patterns evoked by Braille-like dot patterns in the monkey. Journal of Physiology, 310, 117-144.

Katz, D. (1989). The World of Touch (L.E. Krueger, Trans). Hillsdale, N.J.: Lawrence Erlbaum Assoc. (Original work published 1925).

Kirman, J.H. (1974). Tactile apparent motion: the effect of interstimulus onset interval and stimulus duration. Perception and Psychophysics, 15, 1-6.

Kuc, Z. (1989). A Biderectional Vibrotactaile Communication System. PhD thesis, Stanford University, Stanford CA.

Lamb, G.D. (1983a). Tactile discrimination of textured surfaces: Psychophysical performance measurements in humans. Journal of Physiology, 338, 551-565.

Lamb, G.D. (1983b). Tactile discrimination of textured surfaces: neural coding in the monkey. Journal of Physiology, 338, 567-587.

Lederman, S.J. (1981). The perception of surface roughness by active and passive touch. Bulletin of the Psychonomic Society, 18, 253-255.

Lederman, S.J. and Pawluk, D.T. (1993). Lessons from the study of biological touch for robot tactile sensing. In: Advanced Tactile Sensing for Robotics. Ed. H. Nicholls, World Press: London, 1-44.

Loomis, J.M. (1974). Tactile letter recognition under different modes of stimulus presentation. Perception and Psychohpysics, 16, 401-408.

Loomis, J.M. (1980). Interaction of display mode and character size in vibrotactile letter recognition. Bulletin of the Psychonomic Society, 16, 385-387.

Loomis, J.M. (1981). Tactile pattern perception. Perception, 10, 5-27.

Sherrick, C.E. and Rogers, R. (1966). Apparent haptic movement. Perception and Psychophysics, 1, 175-180.

Weisenberger, J.M., & Hasser, C.J. (1994). Role of active and passive movement in vibrotactile pattern percception. Submitted for publication.

White, B.W., Saunders, F.A., Scadden, L., Bach-y-Rita, P., and Collins, C.C. (1970). Seeing with the skin. Perception and Psychophysics, 7, 23-27.

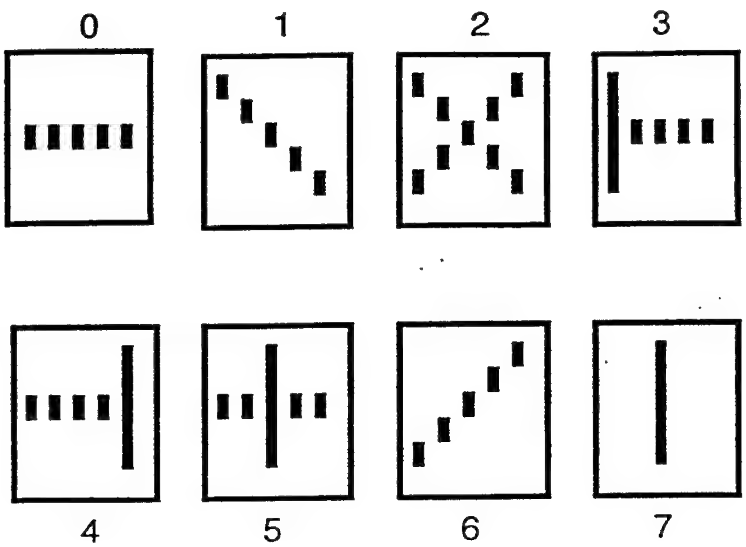


Figure 1a. Icons for stimuli in the set of shapes.

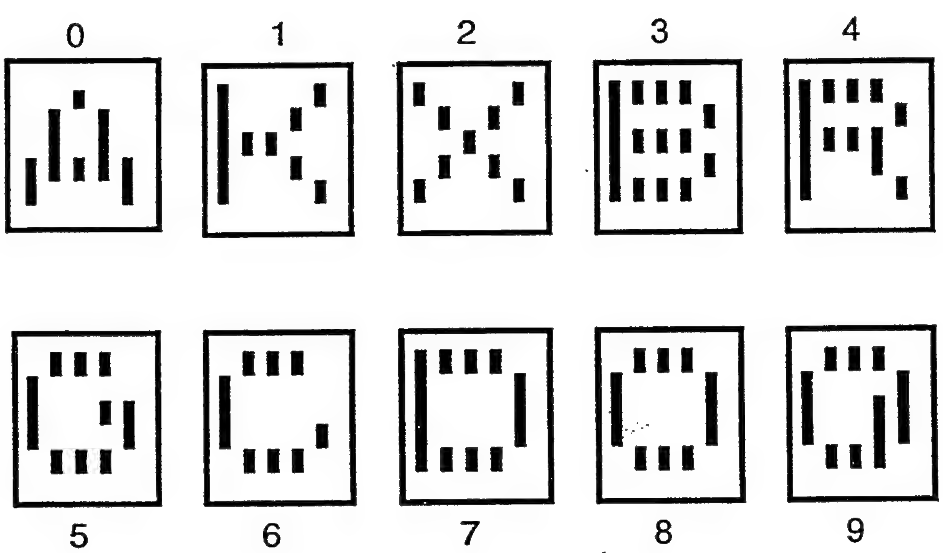


Figure 1b. Icons for the stimuli in the set of letters.

Figure 2. Identification Performance for Displays 1 and 2

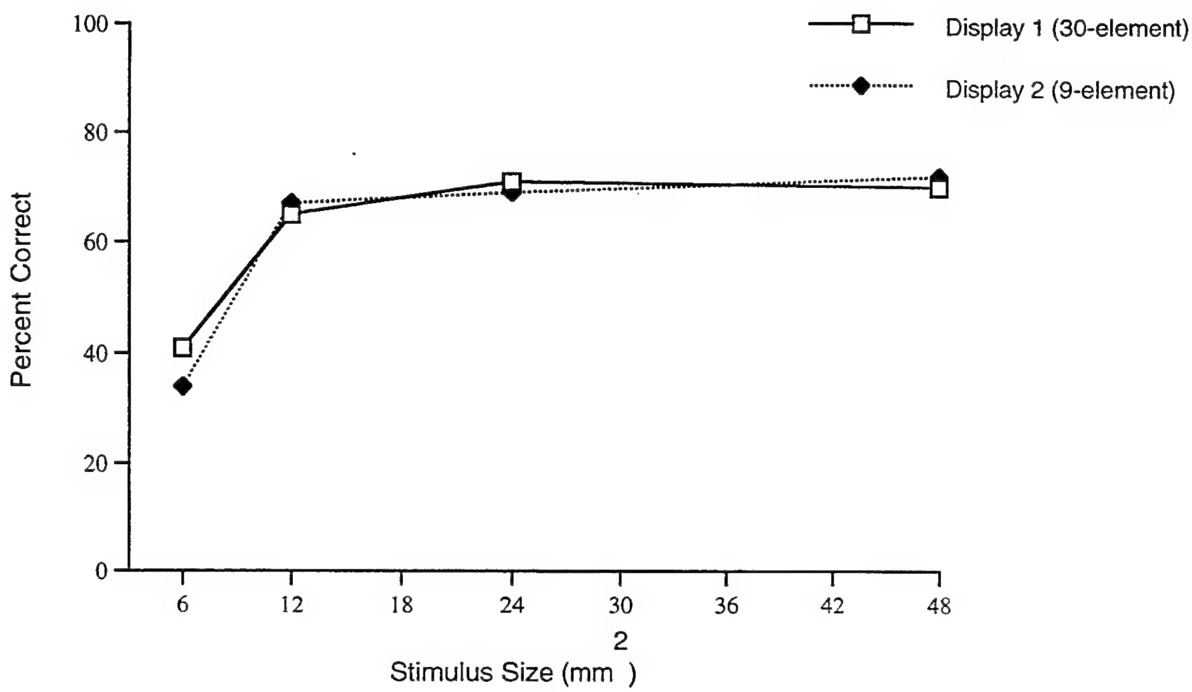


Figure 3. Identification Performance for Displays 3 and 4

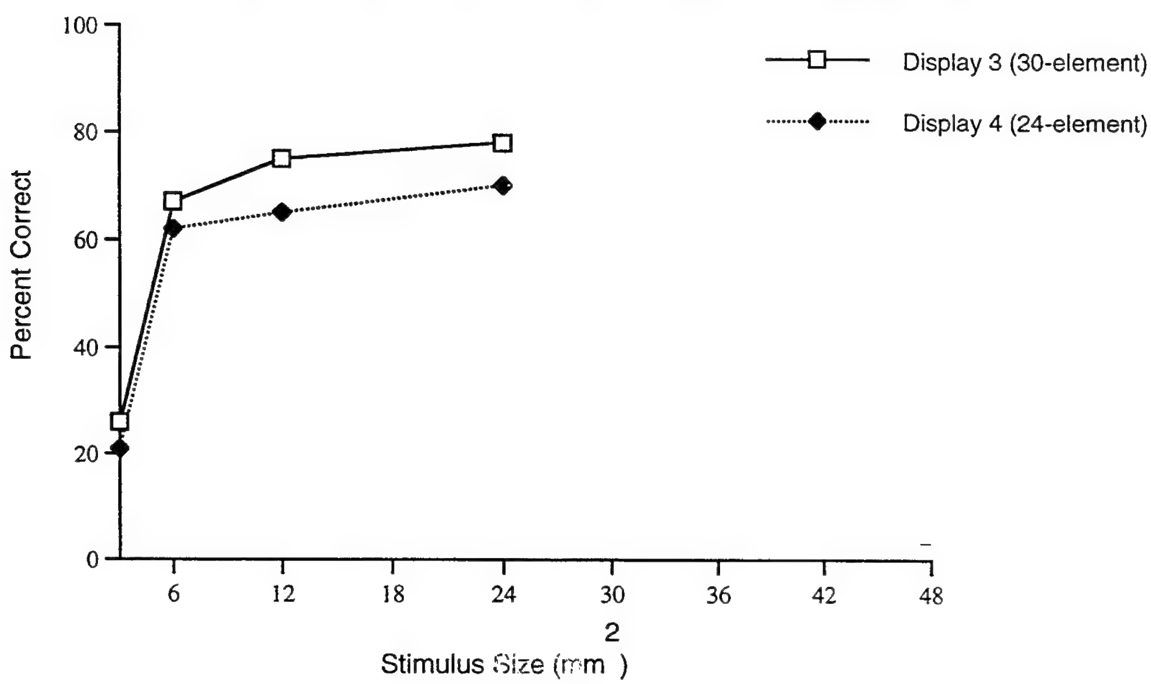


Figure 4. Shape Identification for Different Frequency Pairs

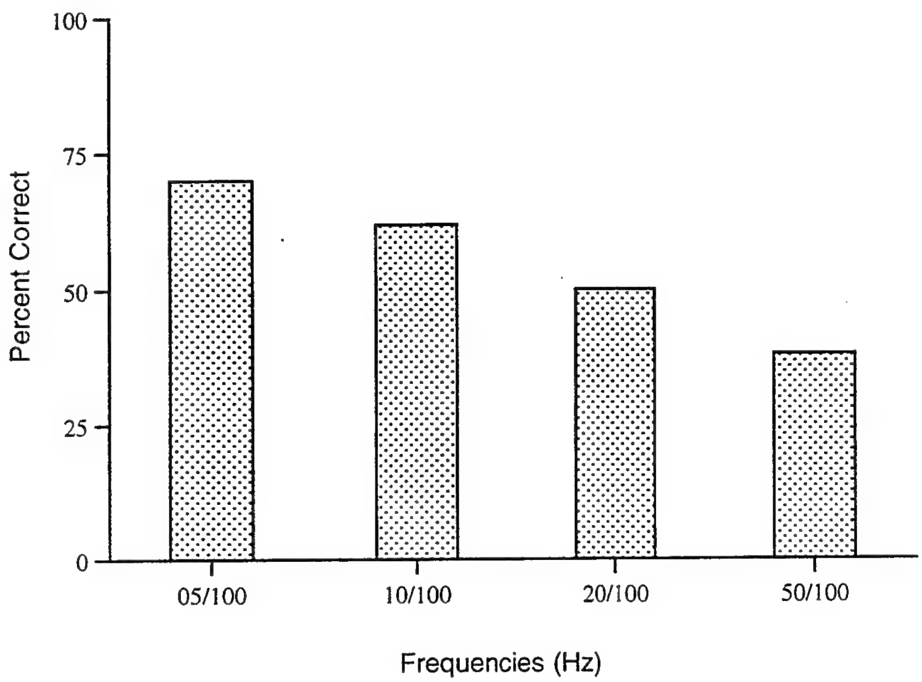
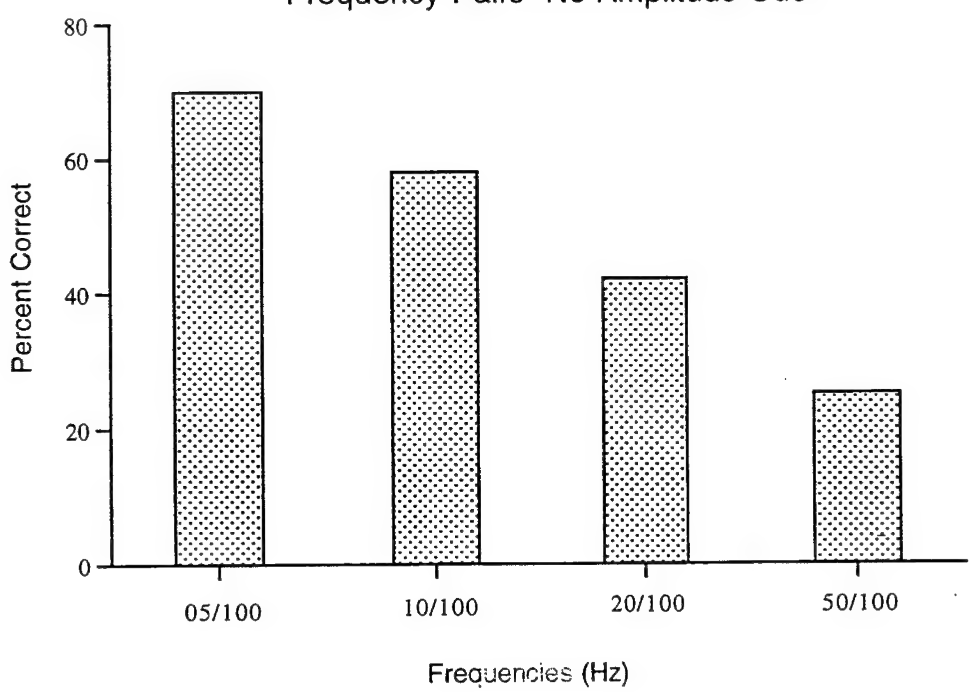


Figure 5. Shape Identification for Different Frequency Pairs--No Amplitude Cue



MELATONIN-INDUCED PROPHYLACTIC SLEEP AS A COUNTERMEASURE FOR SLEEP DEPRIVATION: EVALUATION IN A 36-HR SUSTAINED OPERATION

Rod J. Hughes, Ph.D.
Research Assistant Professor, Department of Psychiatry
Associate Director of Sleep Research, Sleep and Mood Disorders Laboratory

Oregon Health Sciences University
3181 SW Sam Jackson Park Road
Portland, OR 97201

Final Report for:
Summer Research Extension Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Oregon Health Sciences University

February 1996

MELATONIN-INDUCED PROPHYLACTIC SLEEP AS A COUNTERMEASURE FOR SLEEP DEPRIVATION: EVALUATION IN A 36-HR SUSTAINED OPERATION

Rod J. Hughes, Ph.D.
Research Assistant Professor, Department of Psychiatry
Associate Director of Sleep Research, Sleep and Mood Disorders Laboratory
Oregon Health Sciences University

Abstract

The hypnotic efficacy of a single (10 mg) dose of orally administered melatonin given to facilitate sleep in a full-duration daytime episode was assessed in a placebo-controlled, double-blind, cross-over design. Following an eight-hour nighttime sleep opportunity (2300-0700), healthy male (n=3) and female subjects (N=6) were given either melatonin (10 mg) or placebo at 1000 hours. Polygraphically-recorded sleep was assessed in an eight-hour sleep opportunity from 1100 to 1900 hr. Melatonin significantly increased total sleep time, replicating and extending our previous findings. These robust hypnotic effects were reflected primarily in a mean increase in total sleep time of 75 minutes. The sleep induced by melatonin contained significantly more stage 2, less stage 3-4 and a lower percentage of stage 1. In addition to its hypnotic effects, melatonin suppressed core body temperature within 10 minutes of administration, completely suppressing the diurnal rise of core body temperature for at least nine hours. These data clearly demonstrate that a 10 mg dose of melatonin has potent hypnotic and hypothermic effects when administered during the day. The effect of melatonin-induced daytime sleep on nighttime alertness and performance was assessed in 36 hours of sleep deprivation, spanning two nights and one day. This dose of melatonin yielded no carryover sleepiness assessed objectively or subjectively. In fact, melatonin-induced sleep significantly facilitated alertness throughout 36 hours of sleep deprivation. Thus, melatonin appears to be capable of sustaining sleep in the middle of the day leading to increased alertness and performance in subsequent sustained operations.

MELATONIN-INDUCED PROPHYLACTIC SLEEP AS A COUNTERMEASURE FOR SLEEP DEPRIVATION: EVALUATION IN A 36-HR SUSTAINED OPERATION

Rod J. Hughes

Introduction

The pineal hormone melatonin is synthesized and secreted primarily at night in dim light (52, 90). In humans, temporal relationships among melatonin, core body temperature, sleep and alertness begs the question of a direct role for endogenous melatonin in the sleep-wake process. Although a conclusion about endogenous melatonin is premature, administration of exogenous melatonin at nearly all doses has been shown to facilitate sleep onset. We have shown that exogenous melatonin can increase total sleep time in a dose-dependent manner (39, 40).

In seasonally-breeding animals, melatonin is responsible for communicating the circannual changes in photoperiod that determine reproductive timing (e.g., 70). That melatonin also plays a significant role in the circadian system is also well established. In mammals, melatonin serves as a chemical messenger of the primary circadian oscillator, the suprachiasmatic nuclei (SCN), communicating the "darkness" message throughout the body (5, 71). Melatonin acts through high affinity, pharmacologically specific receptors (27, 87). Melatonin binding sites and receptors have been reported peripherally (64) and centrally (for reviews see 46, 57, 79, 80, 82). Specifically, the action of melatonin is primarily through guanine nucleotide binding G protein-coupled receptors. These receptors are found in the SCN (72), where melatonin functions as a true feedback mechanism and where it has direct phase-shifting properties (54).

Melatonin administered to animals (69) and to humans shifts circadian rhythms according to a well-defined phase-response curve (49, 92). Consequently, melatonin administration is efficacious in treating circadian rhythm related sleep disorders. For instance, evening melatonin administration successfully advances sleep and wake times in individuals with delayed sleep phase syndrome (1, 23, 58, 62, 84). Melatonin has also been shown to alleviate jet-lag symptoms after rapid transmeridian travel (3, 18, 61, 66, 67). Similarly, melatonin facilitates adjustment to shiftwork (31, 73, 74). Melatonin can also synchronize free-running circadian rhythms in animals (6) and in humans placed in temporal isolation (56). Consequently, melatonin is effective in entraining free-running sleep-wake rhythms in blind (4, 30, 47, 63, 75, 86) and sighted individuals (55, 83) suffering from Non-24 Hour Sleep-Wake Syndrome. Finally,

anecdotal reports of preliminary clinical trials are reporting that melatonin is quite effective in regulating the sleep-wake cycles of children with various developmental and neurological deficiencies (34, 43). Despite these very promising results, it is not clear whether the effects of melatonin on regulating sleep-wake schedules are a result of its influence on the circadian clock or some direct hypnotic effect.

Since it was first isolated in 1958, tests of high doses of melatonin have yielded demonstrable hypnotic effects (48). These early investigations were often uncontrolled and used very high doses. Research since has shown that exogenous melatonin facilitates sleep onset in diurnal animals (10, 35, 36, 53). In humans, melatonin given at night (after 2200 hr) may facilitate sleep (41, 76, 89) but not always (29, 42, 77), while melatonin given in the evening (before 2130 hr) more reliably shortens sleep onset (21, 60, 93). Daytime administration also consistently facilitates sleep onset latency (2, 19, 20, 26, 39, 40).

Quite simply, administering melatonin during the diurnal portion of the circadian phase, when endogenous levels are very low, induces physiological and behavioral changes that mimic nocturnal functioning (when endogenous melatonin levels are high). Daytime melatonin administration reduces body temperature (13, 14, 17, 25, 33, 37, 39, 40, 51) increases subjective feelings of fatigue and sleepiness (13, 14, 17, 25, 26, 33, 37, 51), impairs performance (2, 25, 26, 51) and shortens sleep onset (2, 19, 20, 26, 39, 40, 60, 85, 88, 93). The doses required to obtain some these effects may be small. Vollrath, et al. (1981) reported reductions of sleep latency with 1.5 mg administered intranasally at 0900 hr. In fact, doses that elevate daytime melatonin levels to at least nighttime physiological levels (e.g., 0.3 mg) reduces oral temperature and shortens sleep onset latencies (26). In short, exogenous melatonin administered out-of-phase with the endogenous nocturnal melatonin rhythm reduces body temperature and shortens sleep latency (see 24, 38 for reviews).

Based upon these findings, some have suggested that low doses of exogenous melatonin are sufficient for facilitating sleep (e.g., 26, 93). Indeed, Dollins and co-workers (1994) reported that doses that raise daytime levels of melatonin to nighttime physiological levels facilitate sleep in a 30 min. nap. In our investigations, however, we have consistently demonstrated dose-response hypnotic and hypothermic effects. For instance, in a placebo-controlled, double-blind crossover design, we administered two doses of exogenous melatonin (10 mg, 100 mg) or placebo to six

healthy males and assessed subjective feelings of alertness. In this investigation, melatonin yielded dose-response effects on sleepiness and oral temperature (see Figure 1 for subjective Fatigue and Vigor). From these data, we initially concluded that melatonin possess dose-response hypnotic and hypothermic properties.

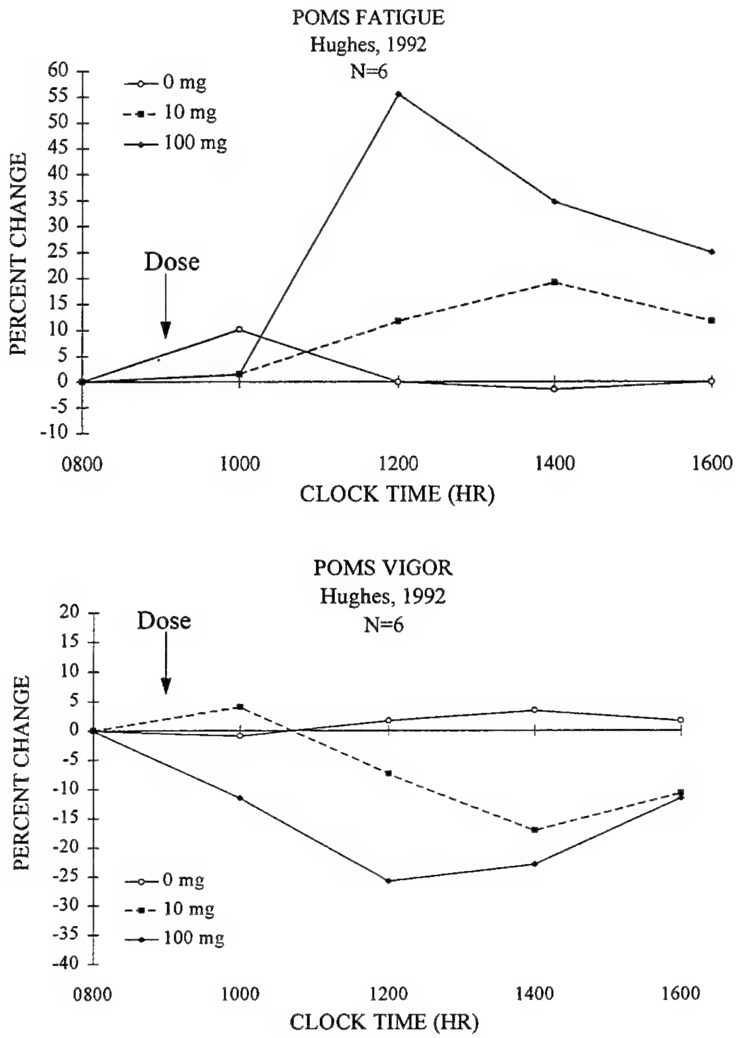


Figure 1. The effects of 10 mg and 100 mg of melatonin on subjective fatigue and vigor (Hughes, 1992).

In another investigation funded by the AFOSR, we recently demonstrated that melatonin administration can significantly facilitate the induction and maintenance of daytime sleep (39, 40). In this investigation, three doses of oral melatonin (1 mg, 10 mg & 40 mg) or placebo at 1000 hr was administered two hours before a moderate-duration daytime sleep opportunity (1200

- 1600 hr). Compared to placebo, melatonin was efficacious in initiating daytime sleep. Melatonin also caused dose-dependent suppression of core body temperature. As depicted in Figure 2, our research shows that melatonin can significantly extend the duration of daytime sleep. Furthermore, melatonin improved daytime sleep without significantly impairing performance tested 7-9 hr after administration (7, 39, 40).

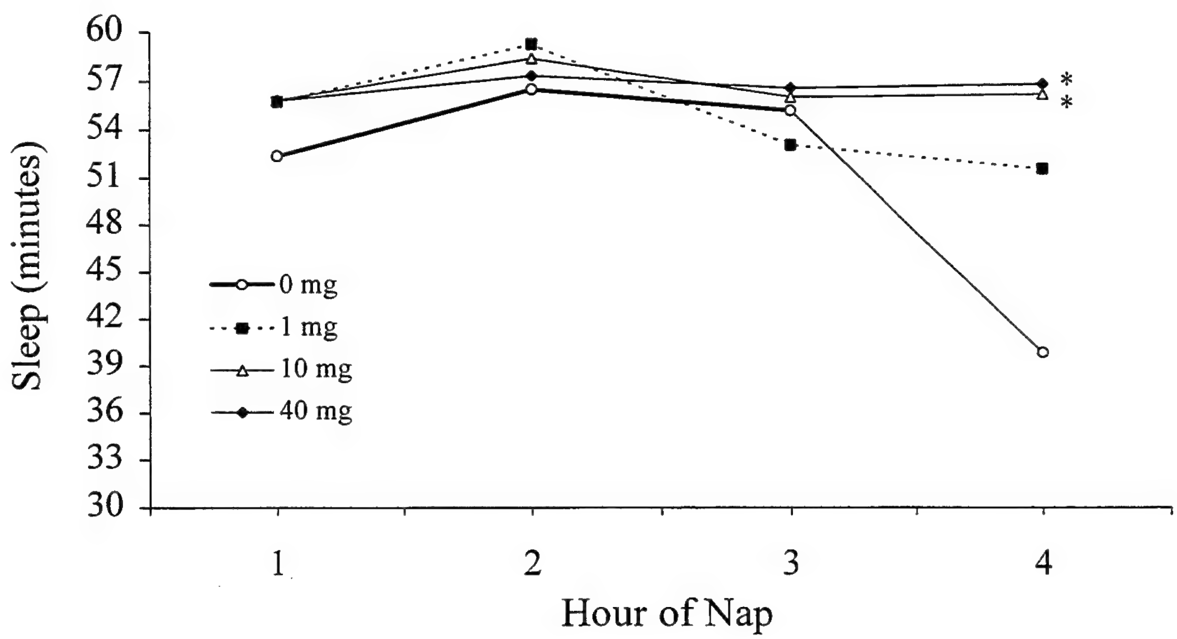


Figure 2. Total sleep per hour of nap for three doses of melatonin and placebo. Oral melatonin was administered at 1000 hr and daytime PSG-recorded sleep episodes were from 1200 hr until 1600 hr. Data indicate a dose-response effect on sleep duration that peaks at 10 mg.

For the first three hours of this nap 1 mg was as efficacious as the higher doses (see Figure 2). However, aside from its effects on sleep latency, the effects of melatonin on sleep duration did not become apparent until two to three hours into the sleep episode, suggesting that doses beyond physiological levels are required to sustain sleep during the day. Thus, although all doses facilitated the <u>induction</u> of daytime sleep, higher doses were required to <u>sustain</u> sleep. This dose-response relationship between melatonin and total sleep time became asymptotic at 10 mg.

In fact, all subjects in the 10 mg and 40 mg conditions were still asleep at the end of the four hour sleep opportunity, suggesting that melatonin could have sustained sleep for longer.

Based upon our previous research, the current investigation was designed to assess the magnitude of the hypnotic and hypothermic effects of exogenous melatonin and to test for nighttime alertness consequences of melatonin-induced daytime sleep. The present investigation employed a placebo-controlled, double-blind, cross-over design to assess the efficacy of a single dose (10 mg) of melatonin in initiating and sustaining a long-duration daytime sleep episode.

The overall goal of the present project was to assess whether melatonin can be used as a natural method to "bank sleep" in preparation for sustained operations or nighttshift-work. The specific goals of the project were to:

Goal #1: Ascertain whether melatonin can sustain a full-duration sleep episode in the middle of the day.

Goal #2: Test the effects of this "pre-sleep" strategy as a countermeasure for sleep deprivation. Thus, following the daytime sleep intervention, performance and sleepiness were assessed in 36 hours of sleep deprivation, spanning two nights.

Methods

Subjects

Twelve healthy male (6) and female (6) volunteers (mean age 25.56 ± 1.56) participated for a total of two sessions each. This report includes analyses of data for the first nine subjects. Most subjects were students at Portland State University. All subjects were screened prior to participation for chronic smoking and heavy alcohol use. Subjects were in good general health, free from major medical disorder and were be able to provide informed consent. Subjects were taking no medications that would affect melatonin production, for example tricyclics or beta blockers. A physical exam was done on all subjects to confirm the absence of a medical disorder. Subjects received $500 upon completion of the investigation.

Dependent Measures

Polysomnographic (PSG) Recordings

Sleep studies were carried out by trained sleep technologists. An initial "diagnostic" polysomnogram (PSG) was performed prior to the first session to rule out primary sleep pathology and to acclimate subjects to sleeping in the CRC. For this sleep study, a standard 16 channel montage was used: four EEG (C_3A_2, O_1A_2, C_4A_1, O_2A_1), four EMG (two for chin, one each for right and left lower extremity), one nasal air flow, two ocular movements, two for respiratory muscle movement, one for oximetry, one for body position and one for EKG. Subsequent sleep recordings omitted the respiratory and leg EMG channels. Sleep records were hand scored using conventional criteria (68). PSG data were collected digitally stored and scored using Sandman® computer hardware and software (Melville Diagnostic Software, Toronto, Canada). Technicians scoring sleep were blind to treatment conditions.

For the entire 36-hr sleep deprivation condition EEG, and EOG were sampled at 256 Hz and stored digitally. Throughout this condition, a modified maintenance of wakefulness test (MMWT) was administered every two hours. The original MWT was modified to include the following: With the lights turned down to less than 20 lux, subjects were asked to lie down in bed and try to remain awake. Because of the funding level of this investigation, we were unable to have attended MWT studies; consequently, subjects were not awakened after sleep onset as in the original MWT. To reduce the amount of cumulative sleep from this modification, the test was reduced from 20 minutes to 10 minutes. Because of this shorter testing time, we choose to modify the MWT further by having subjects begin the test with their eyes closed. This was done to increase the difficulty of staying awake and to shorten the sleep latencies. At the end of the 10 minutes, subjects were instructed to get out of bed and to sit back in the chair.

Core Body Temperature

Core body temperature was assessed using a rectal probe (Mini Logger™, Sunriver, OR). Temperature was recorded every minute throughout the investigation and averaged into ten-minute blocks.

Endogenous Melatonin

All melatonin assays were performed in the Sleep and Mood Disorders Laboratory. Although all of the samples were assayed by RIA, the GCMS assay was used for quality control. The RIA is based on the assay developed by Kennaway (45) and reagents were supplied by ALPCO Ltd. Windham, NH. The lower limit of sensitivity is 1.0 pg/ml; the coefficient of variability is 10.2% for concentrations of 15 pg/ml. The GCMS assay was developed by Lewy and Markey ((50)). The lower limit of sensitivity is 0.5 pg/ml; the coefficient of variability is 2.7 % for concentrations of 20 pg/ml.

Performance Measures

Performance was assessed using three psychomotor and cognitive tasks from the Unified Triservices Cognitive Performance Assessment Battery (UTCPAB). The specific tasks used include: Continuous Recognition task, Procedural Memory reaction time task and the Sternberg Memory Search task (65). To eliminate practice effects, subjects were given training sessions to insure that they were practiced to baseline prior to the first testing day. Performance was tested every four hours during the sleep deprivation phase in the first four subjects and every two hours in the remainder of the subjects tested.

Subjective Ratings

The Stanford Sleepiness Scale is a seven point, self-report, fatigue scale. The scale ranges from 1 (fully alert; wide awake; extremely peppy) to 7 (completely exhausted: unable to function effectively; ready to drop). This scale was administered every hour throughout the sleep deprivation condition. Subjects also completed the Profile of Mood Scales (POMS) every two hours. This particular scale is very sensitive to drug-induced sleepiness (37, 51) and was used primarily to test for carryover sleepiness in the melatonin condition.

<u>Procedure</u>

See Figure 3 for a graphic depiction of the investigation. Subjects were admitted to the Clinical Research Center (from 1700 hr until 0900 hr), prior to the first test session, for a diagnostic circadian phase assessment and for a diagnostic PSG screening. During this first CRC

admission subjects were trained and practiced to baseline on the computer-based performance assessment tasks. An initial "diagnostic" PSG was performed to rule out primary sleep pathology and to acclimate subjects to sleeping in the CRC. For each testing session, subjects returned to the CRC and slept from 2300 hr until 0700 hr. At 0700 hr, subjects were awakened and fed breakfast. Except during sleep episodes and the MMWT, laboratory lighting was set to < 100 lux. At 1000 hr, subjects ingested a pill consisting of either 0 mg or 10 mg of melatonin. Subjects began bio-calibration of the polygraph by 1045 hr. Lights out was at 1100 hr at which time subjects were be instructed to go to sleep. At 1900 hr, lights were turned back up to < 100 lux; subjects then got out of bed and sat in a comfortable chair.

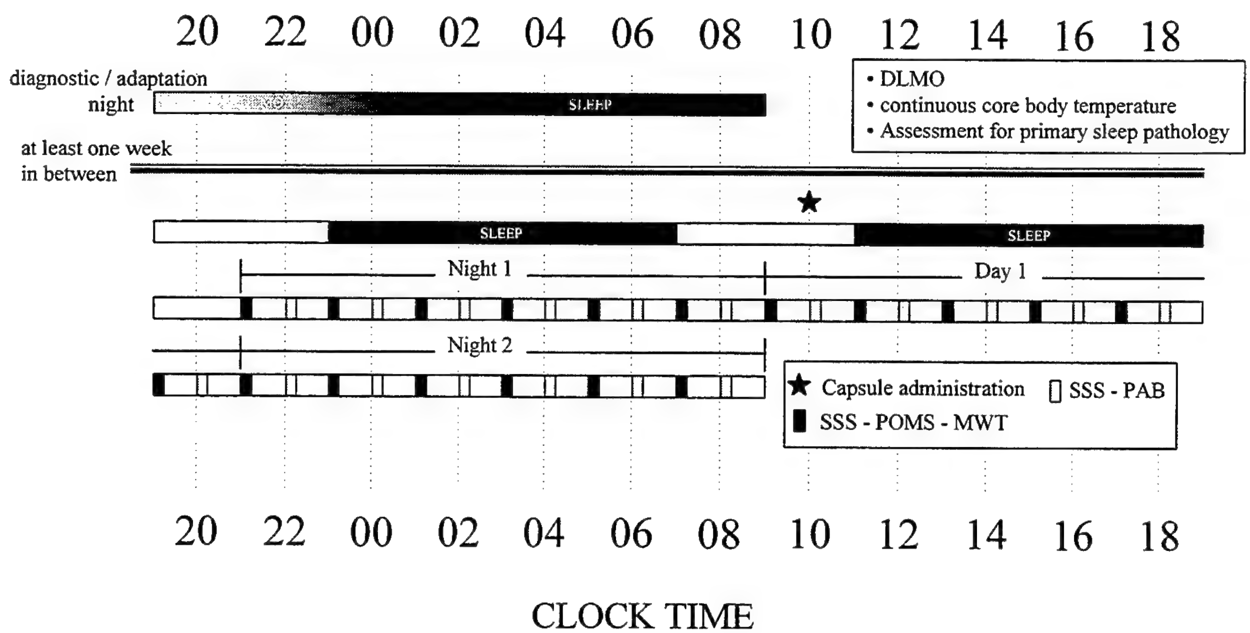


Figure 3. Depiction of experimental protocol. After a full night of sleep, subjects received either 10 mg of immediate release melatonin or placebo at 1000 hr. Daytime sleep was recorded from 1100 until 1900 hr. The effects of this sleep on alertness was then assessed in a 36-hr sleep deprivation protocol.

At 2000 hr, subjects began the 36-hr sleep deprivation phase of the experiment. Performance was assessed every four hours (n=4) or every two hours (n=4). These tests were staggered with the MMWT and the Profile of Mood States (POMS) which were also assessed every two hours. Subjective sleepiness was assessed hourly by the Stanford Sleepiness Scale. When subjects were not taking the PAB or the MWT they were allowed to read, write and watch movies. The testing

session ended after 36 hours (0900 hr at the end of night 2) at which time subjects were released from the study and driven home.

Results

Daytime Sleep

Melatonin increased total sleep time (TST) in all nine subjects. That this increase in daytime TST (Figure 4) was statistically reliable was revealed in a significant one-way repeated-measures analysis of variance (ANOVA) ($F_{(1,8)} = 17.97, p < .003$). Compared to the average total sleep time in the placebo condition (243.12 ± 35.83), melatonin increased total sleep by an average of 75 minutes (317.13 ± 29.10).

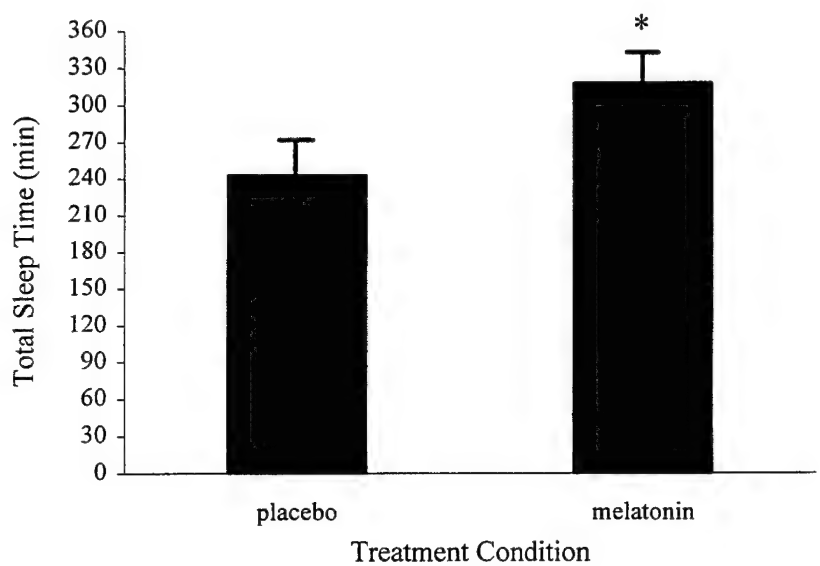


Figure 4. The effect of melatonin on daytime total sleep time. Asterisk represents statistical significance ($p < 0.003$).

Increased sleep in the melatonin condition came primarily from an increase in stage 2 sleep (Figure 5). Thus, sleep following melatonin administration contained an average of 70 minutes more stage 2 sleep ($F_{(1,8)} = 33.26, p < .001$). In addition to more total amounts of stage 2 sleep, melatonin-induced sleep contained a significantly higher percentage of stage 2. Table 1 provides

sleep architecture data presented as percentages of total sleep time. Consistent with better sleep quality, melatonin reduced the percentage of stage 1 sleep. Although melatonin tended to reduce both stage 3 and stage 4 sleep these effects were not statistically reliable. Likewise, the slight trend for melatonin to increase stage REM sleep was not statistically significant.

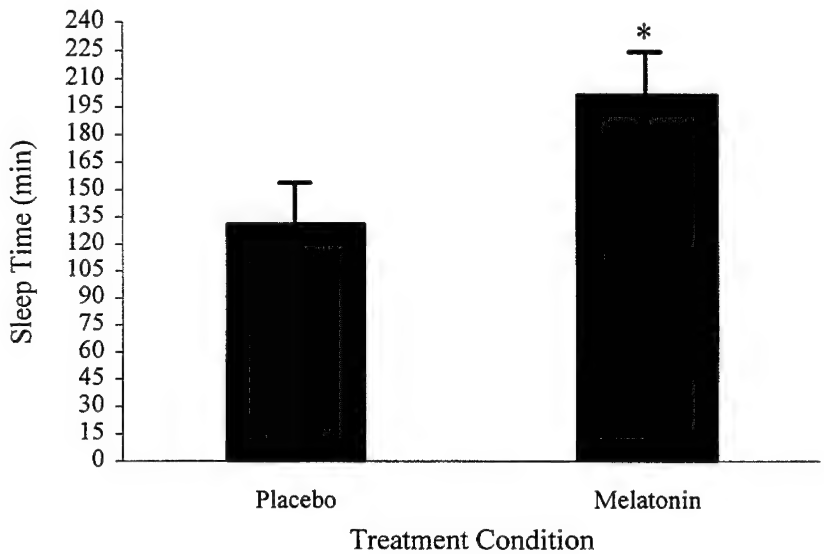


Figure 5. The effect of melatonin on total stage 2 sleep. Asterisk represent statistical significance ($p < 0.001$).

Table 1. Daytime sleep architecture: mean ± (standard error of the mean).

	Melatonin		Placebo		*F* value	*p* value
% Stage 1	25.27	(5.56)	14.92	(1.61)	4.49	0.07
% Stage 2	52.99	(1.58)	63.08	(2.39)	10.47	0.01
% Stage 3	4.93	(1.54)	2.77	(1.38)	1.81	NS
% Stage 4	0.94	(0.49)	0.22	(0.20)	3.03	NS
% Stage REM	15.85	(3.81)	19.02	(1.22)	0.77	NS

Daytime Core Body Temperature

Melatonin suppressed mean core body temperature within 10 minutes of administration (Figure 6). The data are presented and were analyzed as difference scores from core body temperature at the time of capsule administration (1000 hr). Average core body temperature data analyzed in hourly blocks using a two-way repeated measures ANOVA [2 (Treatment) x 9 (Time of day)]. Statistically significant suppression of core body temperature was revealed in a significant main effect for dose ($F_{(1,8)} = 9.78$, $p < .01$). Core body temperature in the placebo condition increased throughout the day, with an average increase of (0.17 C° ± 0.06), while mean core body temperature following melatonin administration did not come back to baseline for at least nine hours; the mean change of core body temperature across the day was negative (-0.08 C° ± 0.08). Thus, melatonin completely suppressed the diurnal rise in core body temperature.

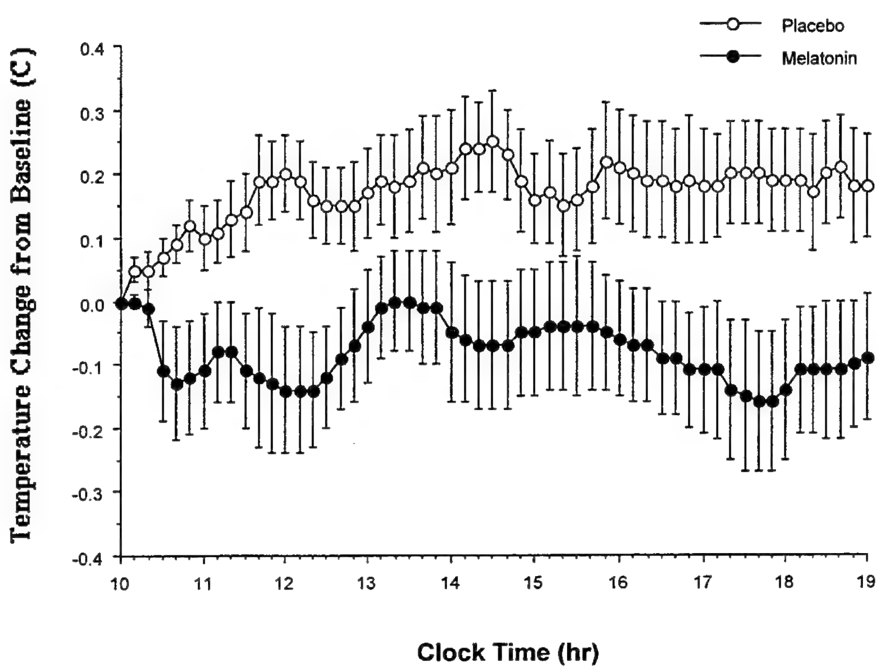


Figure 6. Hypothermic effects of exogenous melatonin administered at 1000 hr. Data are presented as differences from baseline (1000 hr). Enforced sleep times were between 1100 and 1900 hr

Sleep Deprivation Period

Endogenous Melatonin

Figure 7 depicts mean plasma melatonin levels assessed hourly. All 36 hours of data are presented. The figure is divided into night and day for ease of understanding only, as the ambient light intensity was held constant, at less than 100 lux, throughout the sleep deprivation period. Thus, the night and day distinctions do not represent subjective night and day. Plasma melatonin data reveal that exogenous melatonin levels were down to the physiological range by the beginning of the sleep deprivation period (2000 hr) and were below nighttime physiological levels by 2200 hr (twelve hours after administration). The mean melatonin profiles reveal a one hour phase delay of the melatonin rhythm as a result of the intervention.

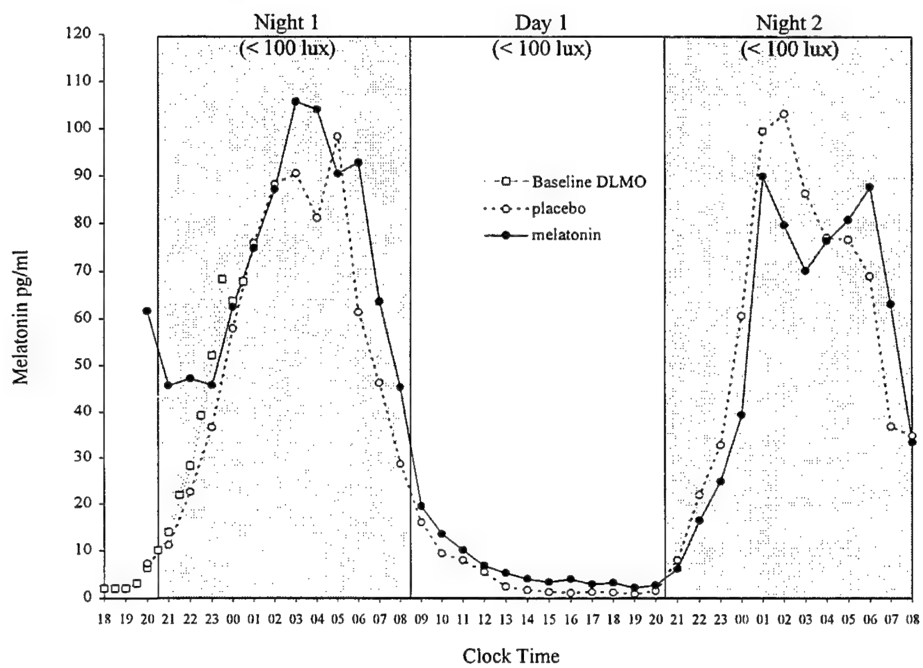


Figure 7. Plasma melatonin levels for placebo (open circles, dashed line) and melatonin (closed circles, solid line). Throughout the 36-hr light was less than 100 lux.

Subjective Sleepiness

Melatonin-induced sleep increased subjective alertness in the sleep deprivation condition. As depicted in Figure 8, subjective sleepiness data revealed that sleep in the melatonin condition was accompanied by a decrease in subjective sleepiness in the second 24 hours of the sleep

deprivation (day 1 and night 2). These data were analyzed by thirds of the sleep deprivation period (night 1, day 1 and night 2). Data for each 12 hour period, were analyzed by a two-way repeated measures ANOVA [2 (Treatment) x 6 (Time)].

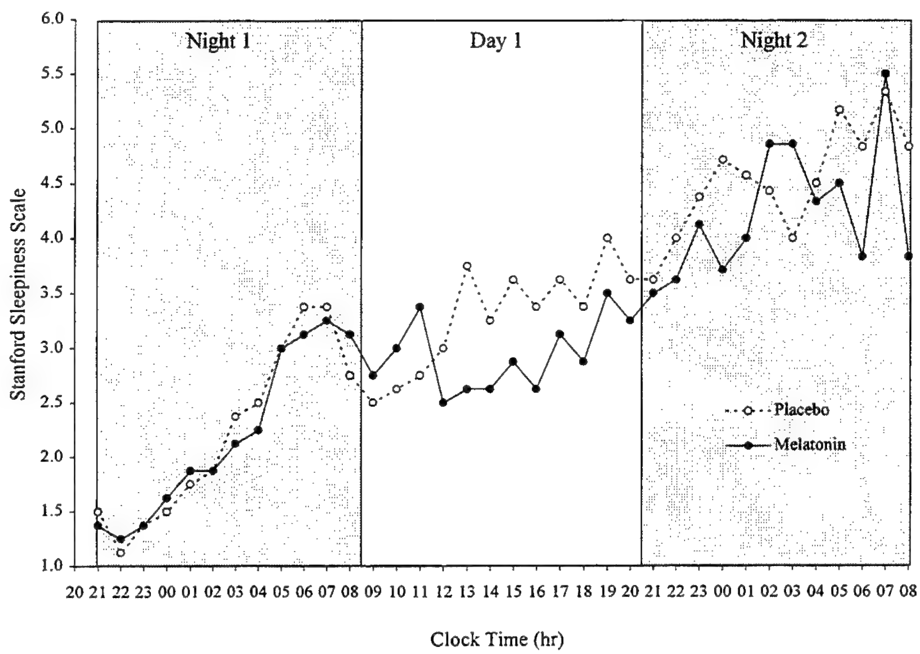


Figure 8. Subjective sleepiness data across 36 hours of sleep deprivation.

Night 1: There were no differences in subjective sleepiness between placebo and melatonin treatment conditions in the first night. As expected, subjective sleepiness increased throughout the night and peaked in the late night and early morning. Thus, sleepiness in night 1 followed the known sleep propensity rhythm. This pattern yielded a significant main effect for Time of night ($F_{(11,77)} = 19.89, p < .0001$). Tukey post-hoc comparisons revealed significant increases in sleepiness in the second half of the night ($p < 0.05$).

Day 1: Differences, in subjective sleepiness, between the melatonin and placebo conditions became apparent in the second 12 hours of sleep deprivation. These differences were reflected in a significant Treatment x Time of day interaction ($F_{(11,77)} = 1.87, p < 0.05$). Post-hoc comparisons revealed that sleepiness in the placebo condition increased significantly across the day so that at 1900 subjects reported feeling significantly more sleepy than at (0900 hr and 1000 hr) ($p <.05$). In the melatonin condition, however, subjects did not report being significantly

more sleepy across this day. Therefore, the additional sleep in the melatonin condition appeared to attenuate the rise in sleepiness mid-way through the sleep deprivation condition.

Night 2: The alerting effects of the melatonin-induced sleep episode extended into the second night of the sleep deprivation condition. Again, differences between the two conditions were reflected in a significant Treatment X Time of day interaction ($F_{(11,77)} = 2.15, p < 0.02$). Although not to the same degree as in day 1, sleep in the melatonin condition suppressed the rise in nighttime sleepiness until 0200 hr. Thus additional sleep in the melatonin condition significantly increased subjective alertness in day 1 and for much of night 2.

POMS Fatigue

As depicted in Figure 9, subjects' self-report ratings of fatigue followed closely that of sleepiness. For this measure, however, the decrease in subjective fatigue associated with additional sleep in the melatonin condition did not achieve statistical significance. The overall main effect for Time of day was significant ($F_{(17,119)} = 9.64, p < 0.0001$). Post-hoc analyses revealed the same form as subjective sleepiness.

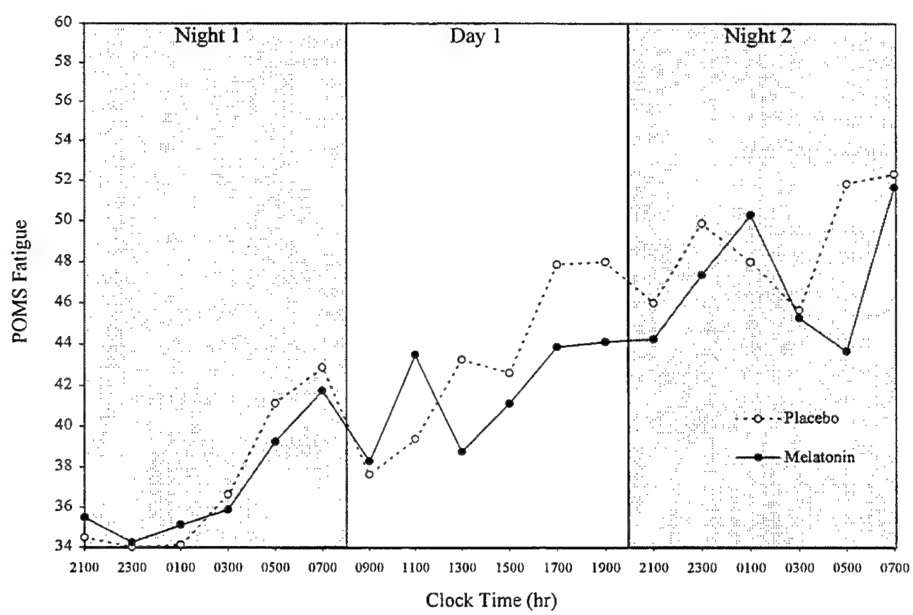


Figure 9. Subjective fatigue data across 36 hours of sleep deprivation.

Modified Maintenance of Wakefulness Test:

Latency to Stage 2 Sleep

Melatonin increased alertness, as measured objectively by sleep latency of the MMWT (Figure 10). Latency to stage 2 sleep data were analyzed in a two-way repeated measures ANOVA [2 (Treatment) x 18 (Time of day)]. That melatonin increased alertness throughout the deprivation condition was revealed by a significant main effect for Treatment condition ($F_{(1,7)} = 6.03$, $p < .05$. Compared to the overall mean latency to stage 2 sleep in the placebo condition (7.49 ± 0.82) subjects took longer to fall asleep in the melatonin condition (8.48 ± 0.66). Increased alertness due to more sleep in the melatonin condition in this objective measure of alertness became apparent in the first night of sleep deprivation.

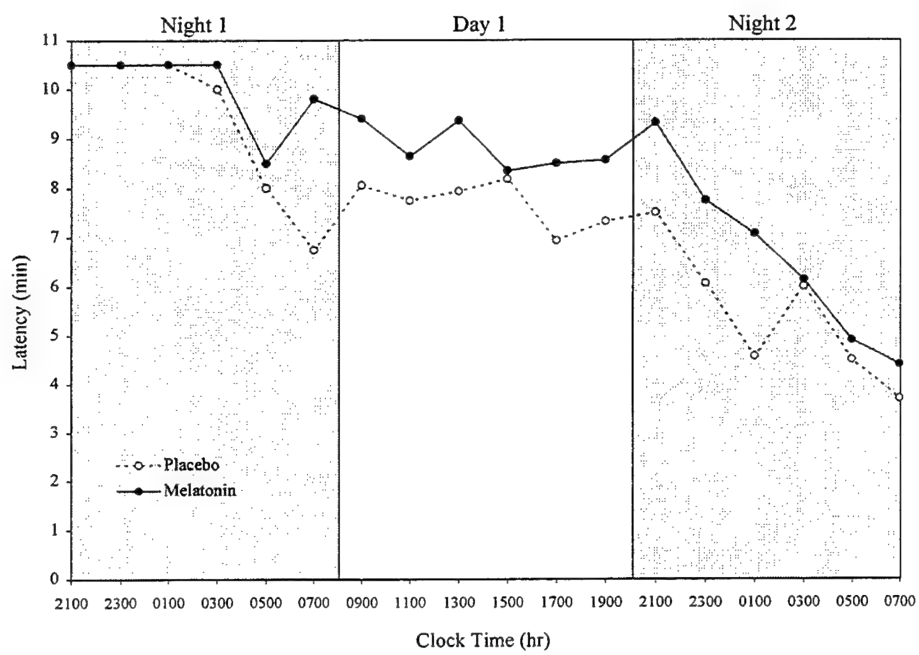


Figure 10. Latency to stage 2 sleep during MMTW. Melatonin consistently increased alertness as measured by this objective test.

Discussion

A 10 mg dose of exogenous melatonin given during the late morning demonstrated robust hypnotic and hypothermic effects. This moderate dose suppressed daytime body temperature within 10 minutes, an effect that lasted at least nine hours. This long-lasting hypothermic effect is consistent with our previous research (37, 39, 40). Melatonin significantly increased the total amount of daytime sleep in an eight-hour daytime sleep opportunity. The form of this sleep was also consistent with our previous research (40, 41), in that melatonin tended to reduce the amount of visually scored slow-wave sleep. This effect, however, was not statistically significant. In this longer duration opportunity melatonin-induced sleep contained a significantly lower percentage of stage 1 sleep, an effect consistent with more restorative sleep. The present investigation supports the hypothesis that the hypnotic effect of exogenous melatonin is at least mediated by its hypothermic effects (9, 38, 40). However, the exact mechanism for the hypnotic effects of exogenous melatonin have yet to be demonstrated.

The increased amount and quality of melatonin-induced daytime sleep opportunities led to increased alertness in 36 hours of sustained operations. Subjective and objectively measured alertness was significantly higher during most of the sleep deprivation condition. Furthermore melatonin yielded no carryover sleepiness with testing beginning one-two hours after the sleep opportunity. This, too, is consistent with our previous findings using performance testing (7). Thus, melatonin at this dose appears to be a good candidate for facilitating daytime sleep; especially in conditions in which individuals must work during their subjective night.

In sustained operations, the fatigue associated with working at night is exacerbated by sleep deprivation. Individuals must cope with both circadian desynchrony and the incremental homeostatic drive for sleep. Under these conditions, individuals become increasingly sleepy and suffer significant performance impairments (15, 44, 78, 91).

Photic countermeasures for sleep deprivation can successfully reverse much of the performance decrements caused by circadian desynchrony (8, 16, 32). Evidence suggests that this is achieved by attenuating the nocturnal rise of endogenous melatonin (8). The application of bright light in the work place, however, is not always possible. For instance, individuals whose jobs require the use of monitors, including AWACs officers and air traffic controllers, are not likely to benefit from this intervention. Pilots flying at night are especially unable to take

advantage of this countermeasure for fatigue. In addition, photic countermeasures do little to alleviate the effects of sleep deprivation.

The use of prophylactic naps have been shown to significantly improve performance in sleep deprivation conditions, suggesting that this intervention may prove an effective countermeasure for sustained operations. Bonnet reported that an 8-hr prophylactic nap opportunity significantly improved sleepiness (MSLTs) and performance for 24 hours of sleep deprivation (11). Bonnet's subjects, however, were not PSG monitored and it is likely that they did not sleep for the entire 8 hours.

The circadian system imposes limits on the duration of sleep initiated out of synchrony with the endogenous sleep propensity rhythm. Daytime sleep is more difficult to initiate, is fragmented by multiple awakenings and is truncated (22, 59, 94). This circadian limitation may prevent daytime naps from completely alleviating prior sleep debt. Further, circadian limitations on daytime sleep may prevent these naps from producing a sleep credit. Daytime sleep limitations may have been responsible for restricting the effectiveness of Bonnet's strategy to only 24 hours of sleep deprivation.

Typical hypnotic drugs, like the benzodiazepines, could be prescribed for facilitating sleep under these circumstances. Temazepam is currently approved for use in Air Force personnel, under supervision of a flight surgeon (e.g., 81). Although Temazepam is effective in promoting and sustaining sleep, it also has negative side effects such as anterograde amnesia (28). Additionally, the relatively long elimination half life of Temazepam (10-17 hr) may impair performance on the first night making it unsuitable for some situations. Benzodiazepine-induced sleep is not as "deep" or as restorative as natural sleep, a fact that could limit their usefulness in this particular intervention. Triazolam, a very short-acting benzodiazepine, has recently been tested for its usefulness in facilitating prophylactic naps prior to sleep deprivation (12). Bonnet reported that Triazolam improved performance over placebo sleep in the second (daytime) half of 24-hr sleep deprivation condition. This drug, however, impaired performance compared to placebo sleep in the first 12 hr of sleep deprivation. Although the short half-life of Triazolam makes it a potential candidate for inducing prophylactic naps, this drug also has the highest report rate for anterograde amnesia (28). The results of the present investigation demonstrate

that melatonin may be superior to traditional hypnotic drugs for facilitating sleep prior to sustained operations.

In conclusion, we have previously shown that melatonin has dose-response hypnotic and hypothermic effects (37, 39), that melatonin can sustain daytime sleep in a four hour sleep opportunity, that this robust hypnotic effect is not accompanied by carryover performance decrements (7) and that in normal subjects a 10 mg dose of melatonin facilitates nighttime sleep as much as a 30 mg dose of Temazepam. The present report supports and extends our previous research on melatonin and sleep. We have now demonstrated the largest treatment effect for exogenous melatonin reported to date. We have also clearly demonstrated that the hypnotic effects of melatonin are not accompanied by carryover sleepiness. Furthermore the present investigation demonstrates increased alertness during 36 hours of sleep deprivation following a melatonin-induced sleep opportunity. Melatonin appears to be an ideal hypnotic for facilitating daytime sleep.

Although these preliminary data suggest that melatonin may be quite useful in facilitating daytime sleep in preparation for sustained or nightshift work, more research is needed before its widespread use can be recommended. This research should include investigations to determine the lowest effective dose for sustaining sleep. For instance, the 10 mg dose may be more effective than lower doses simply because it elevates plasma levels of melatonin above some physiologic or supraphysiologic level for a longer time. If this is so, then a smaller dose delivered in sustained-release formulation may be as effective as the higher dose for sustaining sleep. Melatonin should also be tested as a daytime hypnotic for nightshift workers. Research for this particular indication would take advantage of the unique characteristics of exogenous melatonin as both a hypnotic and a chronobiotic drug.

References

1.Alvarez, B, Dahlitz, MJ, Vignau, J, & Parkes, JD. The delayed sleep phase syndrome: clinical and investigative findings in 14 subjects. *Journal of Neurology, Neurosurgery, and Psychiatry* 1992;55:665-670.

2.Antón-Tay, F, Díaz, JL, & Fernández-Guardiola, A. On the effect of melatonin upon the human brain: it's possible therapeutic implications. *Life Sciences* 1971;10:841-850.

3.Arendt, J, Aldhous, M, & Marks, V. Alleviation of "jet lag" by melatonin: preliminary results of controlled double blind trial. *British Medical Journal* 1986;292:1170.

4.Arendt, J, Aldhous, M, & Wright, J. Synchronisation of a disturbed sleep-wake cycle in a blind man by melatonin treatment [letter]. *Lancet* 1988;i:772-773.

5.Armstrong, SM. Melatonin: the internal zeitgeber of mammals? In: New York: Alan R. Liss, 1989:157-202.

6.Armstrong, SM, Thomas, EMV, & Chesworth, MJ. Melatonin induced phase-shifts of rat circadian rhythms. In: London: John Libbey, 1989:265-290.

7.Badia, P, Hughes, RJ, Wright, KP, Murphy, PJ, & Myers, BL. High Dose Melatonin Facilitates Sleep Without Negative Carryover on Performance. *Sleep Research* 1996;25:In Press.

8.Badia, P, Myers, B, Boecker, M, Culpepper, J, & Harsh, J. Bright light effects on body temperature, alertness, EEG, and behavior. *Physiology and Behavior* 1991;50:583-588.

9.Badia, P, Myers, B, & Murphy, P. Melatonin and thermoregulation. In: Boca Raton: CRC Press, 1992:349-364.

10.Barchas, J, Da Costa, F, & Spector, S. Acute pharmacology of melatonin. *Nature* 1967;214:919-920.

11.Bonnet, MH. The effect of various prophylactic naps on performance, alertness, and mood throughout a 52-hour continuous operation. *Sleep* 1991;14:307-315.

12.Bonnet, MH, & Arand, DL. The effect of prophylactic versus nocturnal naps on performance. *Sleep Research* 1994;23:404.

13.Cagnacci, A, Elliott, JA, & Yen, SSC. Melatonin: a major regulator of the circadian rhythm of core temperature in humans. *Journal of Clinical Endocrinology and Metabolism* 1992;75:447-452.

14.Cagnacci, A, Soldani, R, & Yen, SS. The effect of light on core body temperature is mediated by melatonin in women. *Journal of Clinical Endocrinology and Metabolism* 1993;76:1036-1038.

15.Campbell, SS. Effects of sleep and circadian rhythms on performance. In: New York: Academic Press, 1992:195-216.

16.Campbell, SS, & Dawson, D. Enhancement of nighttime alertness and performance with bright ambient light. *Physiology and Behavior* 1990;48:317-20.

17.Carman, JS, Post, RM, Buswell, R, & Goodwin, FK. Negative effects of melatonin on depression. *American Journal of Psychiatry* 1976;133:1181-1186.

18.Claustrat, B, Brun, J, David, M, Sassolas, G, & Chazot, G. Melatonin and jet lag: confirmatory result using a simplified protocol. *Biological Psychiatry* 1992;32:705-11.

19.Cramer, H, Bohme, W, Kendel, K, & Donnadieu, M. Liberation of growth hormone and melanocyte-stimulating hormone in natural and melatonin-induced human sleep. *Arzneim.-Forsch. (Drug Research)* 1976;26:1076-1078.

20.Cramer, H, Kendel, K, & Beck, U. Influence of melatonin on sleep in humans. In: Basel: Karger, 1972:488-491.
21.Cramer, H, Rudolph, J, Consbruch, U, & Kendel, K. On the effects of melatonin on sleep and behavior in man. In: New York: Raven Press, 1974:187-191.
22.Czeisler, CA, Weitzman, ED, Moore-Ede, MC, Zimmerman, JC, & Knauer, RS. Human sleep: its duration and organization depend on its circadian phase. *Science* 1980;210:1264-1269.
23.Dahlitz, M, Alvarez, B, Vignau, J, English, J, Arendt, J, & Parkes, JD. Delayed sleep phase syndrome response to melatonin. *Lancet* 1991;337:1121-1123.
24.Dawson, D, & Encel, N. Melatonin and sleep in humans. *Journal of Pineal Research* 1993;15:1-12.
25.Dollins, AB, Lynch, HJ, Wurtman, RJ, Deng, MH, Kischka, KU, Gleason, RE, & Lieberman, HR. Effect of pharmacological daytime doses of melatonin on human mood and performance. *Psychopharmacology* 1993;112:490-496.
26.Dollins, AB, Zhdanova, IV, Wurtman, RJ, Lynch, HJ, & Deng, MH. Effect of inducing nocturnal serum melatonin concentrations in daytime on sleep, mood, body temperature, and performance. *Proceedings of the National Academy of Science USA* 1994;91:1824-1828.
27.Dubocovich, ML, & Takahashi, J. Use of 2-(125l] iodomelatonin to characterize melatonin binding sites in chicken retina. *Proc. Natl. Acad. Sci. USA 84* 1987;3916-3920.
28.FDA. *Adverse experiences spontaneous reporting system: comparison of Restoril, Dalmane and Halcione*. Washington DC: US Government Printing Office, 1985.
29.Ferini-Strambi, L, Zucconi, M, Biella, G, Stankov, B, Fraschini, F, Oldani, A, & Smirne, S. Effect of melatonin on sleep microstructure: preliminary results in healthy subjects. *Sleep* 1993;16:744-747.
30.Folkard, S, Arendt, J, Aldhous, M, & Kennett, H. Melatonin stabilises sleep onset time in a blind man without entrainment of cortisol or temperature rhythms. *Neuroscience Letters* 1990;113:193-198.
31.Folkard, S, Arendt, J, & Clark, M. Can melatonin improve shift workers' tolerance of the night shift? Some preliminary findings. *Chronobiology International* 1993;10:315-320.
32.French, J, Hannon, P, & Brainard, GC. Effects of bright illumination on body temperature and human performance. *Annu Rev Chronopharmacol* 1990;7:37-.
33.French, J, Hughes, R, Whitmore, J, Neville, K, Strollo, P, & Reiter, RJ. Diurnal melatonin effects on oral temperature and subjective fatigue. *Sleep Research* 1993;22:83.
34.Gordon, K, Camfield, P, Dooley, J, & Crist, B. Dramatically successful treatment of severe sleep disturbance in developmentally handicapped children with melatonin [abstract]. *Annals of Neurology* 1993;34:504.
35.Hishikawa, Y, Cramer, H, & Kuhlo, W. Natural and melatonin-induced sleep in young chickens -- a behavioral and electrographic study. *Experimental Brain Research* 1969;7:84-94.
36.Holmes, SW, & Sugden, D. Effects of melatonin on sleep and neurochemistry in the rat. *British Journal of Pharmacology* 1982;76:95-101.
37.Hughes, R (1992). *The effects of two doses of exogenous melatonin on temperature and subjective fatigue* . Final report submitted to the Air Force Office of Scientific Research, Bolling, AFB, DC.
38.Hughes, RJ (1994). *Melatonin, body temperature and sleep in humans: A review of a new hypnotic drug* . Final report submitted to the Air Force Office of Scientific Research, Bolling, AFB, DC.

39.Hughes, RJ, Badia, P, French, J, Santiago, L, & Plenzler, S. Melatonin induced changes in body temperature and daytime sleep. *Sleep Research* 1994;23:496.
40.Hughes, RJ, Badia, P, French, J, Wright, K, & Plenzler, S. The effects of exogenous melatonin on body temperature and daytime sleep. *Journal of Sleep Research* 1994;3:111.
41.Hughes, RJ, Sack, RL, Singer, CM, & Lewy, AJ. A comparison of the hypnotic effects of melatonin and temazepam on nocturnal sleep in healthy adults. *Sleep Research* 1995;24a:124.
42.James, SP, Mendelson, WB, Sack, DA, Rosenthal, NE, & Wehr, TA. The effect of melatonin on normal sleep. *Neuropsychopharmacology* 1987;1:41-44.
43.Jan, JE, Espezel, H, & Appleton, RE. The treatment of sleep disorders with melatonin. *Developmental Medicine and Child Neurology* 1994;36:97-107.
44.Johnson, LC, & Naitoh, P (1974). *The operational consequences of sleep deptrivation and sleep deficit* (AGARD-AG-193). Advisory Group for Aerospace Research & Development.
45.Kennaway, DJ, Firth, RG, Philipous, G, Matthews, CD, & Seamark, RF. A specific radioimmunoassay for melatonin in biological tissue and fluids and its validation by gas chromatography-mass spectrometry. *Endocrinology* 1977;101:119-127.
46.Krause, D, & Dubocovich, M. Melatonin receptors. *Annu. Rev. Pharmacol. Toxicol.* 1991;31:549-568.
47.Lapierre, O, & Dumont, M. Melatonin treatment of a non-24-hour sleep-wake cycle in a blind retarded child. *Biological Psychiatry* 1995;38:119-122.
48.Lerner, AB, & Case, JC. Melatonin. *Federal Proceedings* 1960;19:590-592.
49.Lewy, AJ, Ahmed, S, Jackson, JML, & Sack, RL. Melatonin shifts circadian rhythms according to a phase-response curve. *Chronobiology International* 1992;9:380-392.
50.Lewy, AJ, & Markey, SP. Analysis of melatonin in human plasma by gas chromatography negative chemical ionization mass spectrometry. *Science* 1978;201:741-743.
51.Lieberman, HR, Waldhauser, F, Garfield, G, Lynch, HJ, & Wurtman, RJ. Effects of melatonin on human mood and performance. *Brain Research* 1984;323:201-207.
52.Lynch, HJ, Wurtman, RJ, Moskcowitz, MA, Archer, MC, & Ho, MH. Daily rhythm in human urinary melatonin. *Science* 1975;187:169-171.
53.Marczynski, T, Yamaguchi, N, Ling, G, & Grodzinska, L. Sleep induced by the administration of melatonin (5-Methoxy-N-acetyl-trytamine) to the hypothalamus in unrestrained cats. *Experientia* 1964;20:435-437.
54.McArthur, AJ, Gillette, MU, & Prosser, RA. Melatonin directly resets the rat suprachiasmatic circadian clock in vitro. *Brain Research* 1991;565:158-161.
55.McArthur, AJ, Sack, RL, Hughes, RJ, & Lewy, AJ. Melatonin, a stronger zeitgeber than light in some sighted individuals? *Sleep Research* 1995;24a:
56.Middleton, B, Arendt, J, & Stone, BM. Effect of exogenous melatonin on human activity/rest and rectal temperature rhythms and sleep onset times in constant dim light. *Sleep Research* 1995;24a:528.
57.Morgan, PJ, Barrett, P, Howell, E, & Helliwell, R. Melatonin receptors: localization, molecular pharmacology and physiological significance. *Neurochem. Int.* 1994;24:101-146.
58.Nagtegaal, JE, Smits, MG, & Hemmes, AM. Effects of melatonine in a child with delayed sleep phase syndrome. *Sleep-Wake Research in The Netherlands* 1994;5:117-118.
59.Naitoh, P, Kelly, TL, & Englund, C. Health effects of sleep deprivation. In: Philadelphia: Hanley and Belfus, 1990:209-238.

60.Nave, R, Peled, R, & Lavie, P. Melatonin improves evening napping. *European Journal of Pharmacology* 1995;275:213-216.
61.Nickelsen, T, Lang, A, & Bergau, L. The effect of 6-, 9- and 11-hour time shifts on circadian rhythms: adaptation of sleep parameters and hormonal patterns following the intake of melatonin or placebo. In: London: John Libbey, 1991:303-306.
62.Oldani, A, Ferini-Strambi, L, Zucconi, M, Stankov, B, Fraschini, F, & Smirne, S. Melatonin and delayed sleep phase syndrome: ambulatory polygraphic evaluation. *NeuroReport* 1994;6:132-134.
63.Palm, L, Blennow, G, & Wetterberg, L. Correction of non-24-hour sleep/wake cycle by melatonin in a blind retarded boy. *Annals of Neurology* 1991;29:336-339.
64.Pank, SF, Dubocovich, ML, & Brown, GM. Melatonin receptors in peripheral tissues: a new area of melatonin research. *Biological Signals* 1993;2:177-180.
65.Perez, WA, Masline, PJ, Ramsey, EG, & Urban, KE (1987). *Unified Tri-Services Cognitive Performance Assessment Battery: Review and Methodology (U).* (Report No. AAMRL-TR-97-007). Wright-Patterson Air Force Base, OH.
66.Petrie, K, Conaglen, JC, Thompson, L, & Chamberlain, K. Effect of melatonin on jet lag after long haul flights. *British Medical Journal* 1989;298:705-707.
67.Petrie, K, Dawson, AG, Thompson, L, & Brook, R. A double-blind trial of melatonin as a treatment for jet lag in international cabin crew. *Biological Psychiatry* 1993;33:526-30.
68.Rechtschaffen, A, & Kales, A. *A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects*. Bethesda, MD: U.S. Department of Health, Education and Welfare, Public Health Service, 1968.
69.Redman, J, Armstrong, S, & Ng, KT. Free-running activity rhythms in the rat: entrainment by melatonin. *Science* 1983;219:1089-1091.
70.Reiter, R, Rollag, M, Panke, E, & Banks, A. Melatonin: reproductive effects. *Journal of Neural Transmission* 1978;suppl. 13:209-223.
71.Reiter, RJ. Melatonin: the chemical expression of darkness [review]. *Molecular and Cellular Endocrinology* 1991;79:C153-C158.
72.Reppert, SM, Weaver, DR, Rivkees, SA, & Stopa, EG. Putative melatonin receptors are located in a human biological clock. *Science* 1988;242:78-81.
73.Sack, RL, Blood, ML, & Lewy, AJ. Melatonin administration promotes circadian adaptation to night-shift work. *Sleep Research* 1994;23:509.
74.Sack, RL, Blood, ML, & Lewy, AJ. Melatonin administration to night-shift workers: an update. *Sleep Research* 1995;24:539.
75.Sack, RL, Lewy, AJ, Blood, ML, Stevenson, J, & Keith, LD. Melatonin administration to blind people: phase advances and entrainment. *Journal of Biological Rhythms* 1991;6:249-261.
76.Singer, C, McArthur, A, Hughes, R, Sack, R, Kaye, J, & Lewy, A. High dose melatonin administration and sleep in the elderly. *Sleep Research* 1995;24a:In Press.
77.Singer, C, McArthur, A, Hughes, R, Sack, R, Kaye, J, & Lewy, A. Physiologic melatonin administration and sleep in the elderly. *Sleep Research* 1995;24a:In Press.
78.Smith, AP. Time of day and performance. In: New York: Academic Press, 1992:217-235.
79.Stankov, B, Fraschini, F, & Reiter, RJ. Melatonin binding sites in the central nervous system. *Brain Research Reviews* 1991;16:245-256.
80.Stankov, B, & Reiter, RJ. Melatonin receptors: current status, facts and hypotheses. *Life Sciences* 1990;46:971-982.

81.Storm, WG, & Parke, RC. FB-111A aircrew use of temazepam during surge operations. *AGARD-CP* 1986;415:11-12.

82.Sugden, D. Melatonin: binding site characteristics and biochemical and cellular responses. *Neurochem. Int.* 1994;24:147-157.

83.Tomoda, A, Miike, T, Uezono, K, & Kawasaki, T. A school refusal case with biological rhythm disturbance and melatonin therapy. *Brain and Development* 1994;16:71-76.

84.Tzischinsky, O, Dagan, Y, & Lavie, P. The effects of melatonin on the timing of sleep in patients with delayed sleep phase syndrome. In: New York: Elsevier, 1993:351-354.

85.Tzischinsky, O, & Lavie, P. Melatonin possesses time-dependent hypnotic effects. *Sleep* 1994;17:638-645.

86.Tzischinsky, O, Pal, I, Epstein, R, Dagan, Y, & Lavie, P. The importance of timing in melatonin administration in a blind man. *Journal of Pineal Research* 1992;12:105-108.

87.Vanacek, J. Melatonin binding sites. *J. Neurochem.* 1988;51:1436-1440.

88.Vollrath, L, Semm, P, & Gammel, G. Sleep induction by intranasal application of melatonin. In: London: Pergamon Press, 1981:327-329.

89.Waldhauser, F, Saletu, B, & Trinchard-Lugan, I. Sleep laboratory investigations on hypnotic properties of melatonin. *Psychopharmacology* 1990;100:222-226.

90.Weitzman, ED, Weinberg, V, D'Eletto, R, Lynch, H, Wurtman, RJ, Czeisler, C, & Erlich, S. Studies of the 24-hour rhythms of melatonin in man. *Journal of Neural Transmission Supplement* 1978;13:325-.

91.Winget, DM, DeRoshia, CW, Markley, CL, & Holley, DC. A review of human physiological and performance changes associated with desynchronosis of biological rhythms. *Aviation, Space, and Environmental Medicine* 1984;55:1085-1096.

92.Zaidan, R, Geoffrian, M, Brun, J, Taillard, J, Bureau, C, Chazot, G, & Claustrat, B. Melatonin is able to influence its secretion in humans: description of a phase-response curve. *Neuroendocrinology* 1994;60:105-112.

93.Zhdanova, IV, Wurtman, RJ, Lynch, HJ, Ives, J, Matheson, J, Morabito, C, Dollins, AB, Schomer, DL, & Watkins, CJ. Evening administration of melatonin promotes sleep in humans. *Soc. Neurosci. Abs.* 1994;20:1440.

94.Zulley, J, Wever, R, & Aschoff, J. The dependence of onset and duration of sleep on the circadian rhythm of rectal temperature. *European Journal of Physiology* 1981;391:314-318.